

GenCore version 5.1.6
(C) 1993 - 2004 Compugen Ltd.

Vi protein - protein search, using sw model

On: January 30, 2004, 13:12:38 ; Search time 22 Seconds
2313.635 Million cell updates/sec
(without alignments)

Title: US-09-769-699-2
Effect score: 6294
Sequence: MENTQKTVTVPTGFLGVVA.....DEFLDSGIPIKHGNITMEM 1203

Coring table: BLOSUM62

Gapext: 0.5

Searched: 328717 seqs., 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:
1: /cgm2_6_ptodata/1/1aa/5A_COMB.pep: *
2: /cgm2_6_ptodata/1/1aa/5B_COMB.pep: *
3: /cgm2_6_ptodata/1/1aa/5A_COMB.pep: *
4: /cgm2_6_ptodata/1/1aa/5B_COMB.pep: *
5: /cgm2_6_ptodata/1/1aa/PCUS_COMB.pep: *
6: /cgm2_6_ptodata/1/1aa/backfiles.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Length	Match Length	DB ID	Description
1	3054	48.5	1452	4	US-09-127-227-2
2	183	2.9	35	3	US-09-127-33B-9
3	183	2.9	35	3	US-09-138-876-9
4	131	2.1	757	3	US-09-413-514-84
5	123	2.0	1074	4	US-09-071-035-358
6	123	2.0	1074	4	US-09-071-035-394
7	118.5	1.9	370	4	US-09-352-991A-16913
8	106.5	1.8	5215	3	US-09-105-537-2
9	106.5	1.7	3567	2	US-07-652-734-C-4
10	106.5	1.7	3567	3	US-08-039-009A-4
11	106	1.7	774	1	US-07-31-157-A-7
12	106	1.7	774	2	US-08-541-780-7
13	105.5	1.7	1095	4	US-09-107-532A-3855
14	105	1.7	774	1	US-08-633-760-46
15	105	1.7	774	1	US-08-633-760-48
16	104.5	1.7	635	4	US-08-931-608A-5
17	104.5	1.7	774	1	US-08-019-870-5
18	103	1.6	774	1	US-08-019-870-3
19	103	1.6	774	1	US-08-633-760-44
20	102	1.6	729	1	US-08-070-16F-6
21	102	1.6	729	2	US-08-885-418-6
22	102	1.6	774	1	US-07-935-312-3
23	102	1.6	774	1	US-07-935-312-3
24	102	1.6	774	1	US-08-633-760-50
25	101.5	1.6	773	1	US-08-019-870-1
26	101.5	1.6	773	1	US-08-019-870-1
27	101.5	1.6	774	1	US-08-019-870-8

result No.	Score	Query Length	Match Length	DB ID	Description
1	3054	48.5	1452	4	US-09-127-227-2
2	183	2.9	35	3	US-09-127-33B-9
3	183	2.9	35	3	US-09-138-876-9
4	131	2.1	757	3	US-09-413-514-84
5	123	2.0	1074	4	US-09-071-035-358
6	123	2.0	1074	4	US-09-071-035-394
7	118.5	1.9	370	4	US-09-352-991A-16913
8	106.5	1.8	5215	3	US-09-105-537-2
9	106.5	1.7	3567	2	US-07-652-734-C-4
10	106.5	1.7	3567	3	US-08-039-009A-4
11	106	1.7	774	1	US-07-31-157-A-7
12	106	1.7	774	2	US-08-541-780-7
13	105.5	1.7	1095	4	US-09-107-532A-3855
14	105	1.7	774	1	US-08-633-760-46
15	105	1.7	774	1	US-08-633-760-48
16	104.5	1.7	635	4	US-08-931-608A-5
17	104.5	1.7	774	1	US-08-019-870-5
18	103	1.6	774	1	US-08-019-870-3
19	103	1.6	774	1	US-08-633-760-44
20	102	1.6	729	1	US-08-070-16F-6
21	102	1.6	729	2	US-08-885-418-6
22	102	1.6	774	1	US-07-935-312-3
23	102	1.6	774	1	US-07-935-312-3
24	102	1.6	774	1	US-08-633-760-50
25	101.5	1.6	773	1	US-08-019-870-1
26	101.5	1.6	773	1	US-08-019-870-1
27	101.5	1.6	774	1	US-08-019-870-8

RESULT 1
US-09-127-227-2
; Sequence 2, Application US/09127227
; Patent No. 639934
; GENERAL INFORMATION:
; APPLICANT: David M. Knipe
; APPLICANT: Elizabeth McNamee
; TITLE OF INVENTION: Replication-Competent Virus Expressing A Protein
; FILE REFERENCE: HU98-05
; CURRENT APPLICATION NUMBER: US/09/127,227
; CURRENT FILING DATE: 1998-07-31
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: herpesvirus
US-09-127-227-2

Query Match 609; Conservative 200; Mismatches 366; Indels 48; Gaps 17;

Best Local Similarity 49.8%; Fred. No. 2.7e-304; Score 3054; DB 4; Length 1452;

Query 1 MENTOKTVT--VPTGLPCTYV--ACRVEDDLETSFLALARSTSDALLPLMLNLTVE 55
Db 1 METPKPKATTIVKPPGLGTYVACPSEGEL--LALLSARSDDADVAVPLVYGLTVE 58
Query 56 KTFSSLLAVSVGARTGLAGAGITLKLTSHFYPSPVFGKHYLPSSAAPNLTRACNA 115
Db 59 SGEEANVAVSVRSRTISLGGAWSLKLTPSWSSTSYVFGGRHLDPSTQPNLRLCR 118
Query 116 ARERFGFCSRQQPPDVCAEVTTGAEITCRUGLEPENTILYLVVTAFLKEAVFMCNVFLH 175
Db 119 ARHFGFEDYTPPGDQKHTGEALRGLDPEALLVVTEKFKEACVINTFLH 178
Query 176 GGDIVHNHGDIVTRIPFLPQVLFVNRLVPDPNTNHRSIGEGFVPPFYNTGLCH 235
Db 179 GGSKVTKTGGAYVRIPYPLQFLMPDFSRVIAEPNNHSIGENFTPPFFNPLNR 238
Query 236 LIHDICVIAFMVALVRVNTAVARGAAHLFDENHECAVLDIDITYFQSSSSGTTAR 29
Db 239 LLIFAVVGPAAVLRCRNVDAYAAAHLFENHEGAALPDITTAFAESLSQG--KTPR 29
Query 296 GARNDVNSTSKEPSGGERLAISMAADLHAEVFNTRGIEEPTDKWPMFIGM 355
Db 297 GGR ---DGCGKCPAGGFQORLAVNAGDAALALESVSVAWFDEPDT-SAWPLCEQQ 351
Query 356 EGTLPRLNALGSYTARVAGVIGAMVESPNSALYLTVEDSLGMTEAKDGGPFSNRPYQF 415

SEARCH 2 21/04
NOTES 09/769,699

RESULT 2
US-08-807-332B-9
Sequence 9, Application US/08807332B
; Patent No. 598074
GENERAL INFORMATION:
; APPLICANT: Dreyfus, David H.
; ATTORNEY/AGENT INFORMATION:
; TITLE OF INVENTION: PRODUCTS AND PROCESSES FOR REGULATION OF
; NUMBER OF SEQUENCES: 32
; ADDRESSEES: Sheridan Ross
; STREET: 1700 Lincoln St., Suite 3500

Db 352 Draaaravgtalaraxglyanfestnsalhetevdagpadpkdhsk-pspyrefly 410
Qy 416 Aghilaanpqtdrdghvl-----ssostgsntefsvdylallicfgaflarlfyl 468
Db 411 Pgthvaanpqdreghvvpgfegraptalvlgcq-efagehmilcfspsalakmlfyl 469
Qy 469 Ercdagafatgggh-gdalxytgtpdsepcslcbktrpvcaahtrverlormprgqat 527
Db 470 Erddggvgromdvervqadsnotdvcncltedfirhacttlmrlrphpkasaa 529
Qy 528 Ropigvftgmsoyscdplgnayplirkpgdoteakatmdtyratlerlfidleq 587
Db 530 Rgaiqvgftgmmsmyscdvlgnrafaflkrr-adgsetpartimqetraatervmarelet 588
Qy 588 Erldrgapesssegssivvvhdpfrlltdflarierottomkvkvetrykriegls 647
Db 589 Lqyvqavtamggleit1tnrealhtvnnntqvvdrevemnlvegrnkfrpdglg 648
Qy 648 Bathsmalifdpysgafcpitnfvkrthlavtdlaisochyfyqcovegrnfrnqfo 707
Db 649 Eniamslstldpacygcpcplqlqrgrenlavqfdlalsochvfaaggsvbervfrnqfo 708
Qy 708 Pvllerfrfdlengbrisritvltseg-pvsapnptiqodafagrfedgdilarvsevi 766
Db 709 Pvlerrvdmnfngflsaktltvalsegaaicapsltagotaesffegdyarvtlgefp 768
Qy 767 Rdtrkgnrufvsgntnlstaaarlvglasatorqerdrvdmngflcoffgllfp 826
Db 769 Kelvksrsvlafgasanaeakarvalosavokpdrvdvllglgfllkfqaiffp 828
Qy 827 Rgmpnskspnponewtllorompadktltheittiaavkrteeyaaainfnlpptci 886
Db 829 Ngkpgsnqnqnpnwtaqlornqdlparlsredetifikkslidgaintlnanrv 988
Qy 887 Gelaofymanlikycdhsolylintltgairprpssvleirwdtsadiotqa 946
Db 889 Selamymanolikycdhsyfintlatatagsrrpsvqaaw--saqggagleaga 945
Qy 947 Kalltektenpelwttatfstthlyarnqrpmtygdiskyhagnnrvqagmwsg 1006
Db 946 Ralm davdaahpgantsmfascnlrrpmaarpvylglisikygmndrvqagnwas 1005
Qy 1007 Lnggnrvnvcpletdtrtrfliaqrggfcpcvpgssgnretlisdqvrqtiisggamq 1066
Db 1006 Imggnaacplifdrkrkeflacpragfcaasnlggahessicbeorglseggava 1065
Qy 1067 Laaytvvragarohmafdwlsttddeflardlehdqliotltpwtvbgal--- 1123
Db 1066 Ssvfatvkslsgprtqolqedwalleylemmeltaralerngewstdaaleva 1125
Qy 1124 -Eavkldekitagdgetptnlafnfd--scpshtdttsnvnlnsgsnsstvpglkrp 1180
Db 1126 Heaealvsqlnqage----venfgdfge---ddnatppfgpgapgapafgrkra 1173
Qy 1181 Pdddefldlsqipkhgnitmem 1203
Db 1174 Fhddppg-egpdkgdltlum 1195

RESULT 3
US-09-338-876-9
Sequence 9, Application US/09338876
; Patent No. 6138584
GENERAL INFORMATION:
; APPLICANT: Dreyfus, David H.
; ATTORNEY/AGENT INFORMATION:
; TITLE OF INVENTION: GENE RECOMBINATION
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS: 32
; ADDRESSEES: Sheridan Ross
; STREET: 1700 Lincoln St., Suite 3500
; CITY: Denver
; STATE: CO
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/Ms-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/091338, 876
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/807/332
; FILING DATE: 28-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kovari, Joseph E.
; REGISTRATION NUMBER: 33,005
; REFERENCE DOCUMENT NUMBER: 2879-39
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/863-9700

Db 588 Erldrgapesssegssivvvhdpfrlltdflarierottomkvkvetrykriegls 647
Qy 589 Lqyvqavtamggleit1tnrealhtvnnntqvvdrevemnlvegrnkfrpdglg 648
Db 648 Bathsmalifdpysgafcpitnfvkrthlavtdlaisochyfyqcovegrnfrnqfo 707
Qy 649 Eniamslstldpacygcpcplqlqrgrenlavqfdlalsochvfaaggsvbervfrnqfo 708
Db 709 Pvlerrvdmnfngflsaktltvalsegaaicapsltagotaesffegdyarvtlgefp 768
Qy 767 Rdtrkgnrufvsgntnlstaaarlvglasatorqerdrvdmngflcoffgllfp 826
Db 769 Kelvksrsvlafgasanaeakarvalosavokpdrvdvllglgfllkfqaiffp 828
Qy 827 Rgmpnskspnponewtllorompadktltheittiaavkrteeyaaainfnlpptci 886
Db 829 Ngkpgsnqnqnpnwtaqlornqdlparlsredetifikkslidgaintlnanrv 988
Qy 887 Gelaofymanlikycdhsolylintltgairprpssvleirwdtsadiotqa 946
Db 889 Selamymanolikycdhsyfintlatatagsrrpsvqaaw--saqggagleaga 945
Qy 947 Kalltektenpelwttatfstthlyarnqrpmtygdiskyhagnnrvqagmwsg 1006
Db 946 Ralm davdaahpgantsmfascnlrrpmaarpvylglisikygmndrvqagnwas 1005
Qy 1007 Lnggnrvnvcpletdtrtrfliaqrggfcpcvpgssgnretlisdqvrqtiisggamq 1066
Db 1006 Imggnaacplifdrkrkeflacpragfcaasnlggahessicbeorglseggava 1065
Qy 1067 Laaytvvragarohmafdwlsttddeflardlehdqliotltpwtvbgal--- 1123
Db 1066 Ssvfatvkslsgprtqolqedwalleylemmeltaralerngewstdaaleva 1125
Qy 1124 -Eavkldekitagdgetptnlafnfd--scpshtdttsnvnlnsgsnsstvpglkrp 1180
Db 1126 Heaealvsqlnqage----venfgdfge---ddnatppfgpgapgapafgrkra 1173
Qy 1181 Pdddefldlsqipkhgnitmem 1203
Db 1174 Fhddppg-egpdkgdltlum 1195

TELEFAX: 303/863-0223 SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide

Query Match 2.9%; Score 183; DB 3; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.e-11;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 545 DPLGNAYAPYLIRKPGDQTAAKATMQDTYRATLE 579
Db 1 DPLGNAYAPYLIRKPGDQTAAKATMQDTYRATLE 35

RESULT 4
Sequence 84, Application US/09413814
Patent No. 6225065
GENERAL INFORMATION:
APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
APPLICANT: Bristol-Myers Squibb, Co.
APPLICANT: Beyer, Stefan
APPLICANT: Bloecker, Helmut
APPLICANT: Brandt, Peirra
APPLICANT: Cino, Paul M
APPLICANT: Dougherty, Brian A
APPLICANT: Goldberg, Steven L
APPLICANT: Mueller, Joachim
APPLICANT: Reichenbach, Hans
TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
TITLE OF INVENTION: heteropolyketide compounds
FILE REFERENCE: PCT/US 99/2355
CURRENT FILING DATE: 1999-10-07
EARLIER APPLICATION NUMBER: US/09/413,814
EARLIER FILING DATE: 1998-10-09
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 84
LENGTH: 757
TYPE: PRT
ORGANISM: Sorangium cellulosum

Query Match 2.1%; Score 131; DB 3; Length 757;
Best Local Similarity 21.1%; Pred. No. 0.001; Mismatches 251; Indels 212; Gaps 37;

Db 168 ----AALQRLRMIAATDPS-POPIEGFGA---AIRPESV----- 198

Qy 54 VENTFTSSLAVVSGART---TGLAGAGITKLTTSHFYPVFHGGKHLPPSSAAPNL 109
Db 129 LERTLPRVASIVDARAEEAATLSAVAGIVAS-----LPASAAA-- 167

Qy 110 TRACNAARERFGSRCQQPPVNDGDAVETGAGTCIGLEPEINTLYVLTAFRAVFM 169
Db 168 ----AALQRLRMIAATDPS-POPIEGFGA---AIRPESV----- 198

Qy 170 NPYFLHY----GGIDIVHNGDGVIRPLPFVQLFMDVNRLVPDPFNTHRSIGRGFVY 224
Db 199 -AFQYTGSTGFBPKGMNLTHGL---LHSNLLIAFGFDLTSFDPV-----GVIV 244

Qy 225 PTPhYNTGLC----HLIHDVIAFMVALR---VRNNTAV---ARGAAHIAFD-- 267
Db 245 LPP4HDMGLIGGIQALYRRIRVALMSLPSLFLQRPMWLRAYSAGVSQGPNFAYDLC 304

Qy 268 ---ENHEAVLLEPDITYFQSSSGTTTANGARRNDVNTSKPSGGERLASIMA 323
Db 305 VRKSSEEEERAALD-----LREWEVAPTGAEEFVRADTLDRPARAPAVSGFRRE--AFYP 355

324 ADTALHAEVIFNTGTYEETPTDIKEWPMEFIGMGTIPRNALGSYTAvgVIGAMVF-- 381
Db 356 CYGLATLIVSGGARAEAFV-----LARLAPBEVELLARAVASARE--GARYFVG 403

Qy 382 ----SPNSALYLTVEDSGMTEAKDGGPGPSFRYQFAGPHIA---ANPQTDRDGHV 432
Db 404 SGRAIDPRA---VAIVDPAG---NELGPFG-EIGEWV-VSGPSVAVGTMWRPE----- 447

Qy 433 ISSQSTGSSNNTEFSVDYLALICGGAPlLARLFLYLEDGAGTGGHGDALKYVTGT-- 490
Db 448 -----ETEAETGATLAGSAAAPR----YLRTGDLGFLRGGE---LFVVGRSK 486
Qy 491 -----FDSEIIFCSLOCBKH---TRPVYC-AHHTVHBLRQRM-----PFFGQATRQ 529
Db 487 DLIILRGRNHPDQEKTVESSHRAVPGCGSARSVEHGERLAVVCEVDPPVAAADRE 546
Qy 530 PIGVFCGMNSOYSDCDPLGNAYAPYLIRKPGDQTEAAKATMQDTYRATLERLFDLDEQR 589
Db 547 IVAAREAVTAHQ---LVAHAVALLIAPGALPKTSSGKVRRRECRAFIE-----DALGER 598

Qy 590 -----LUDGAPCSSEGSLSVVDPHPTFRFLDTRLARIEOTTTQFMKVLYBTRDKI 642
Db 599 HVAFAPFLDDASPPD--APPETBEPSPRSLSLALRS---TLARALRLDAGOIDAL 651

RESULT 5
US-09-071-035-358
Sequence 358, Application US/09071035
Patent No. 6418043
GENERAL INFORMATION:
APPLICANT: Gil H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION NUMBER: US/09/071,035
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
PRIORITY NUMBER:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 35,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 358:
SEQUENCE CHARACTERISTICS:
REF ID: 358
SEQUENCE:
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-09-071-035-358

Query Match Score 123; DB 4; Length 1074;
 Best Local Similarity 18.3%; Prod. No. 0.013;
 Matches 206; Conservative 150; Mismatches 415; Indels 352; Gaps 52;

250 RVNNTAVARGAAHLAFAFDENHEGAVLPPDITYTFOSSSSGTTARGARRNDVNSTSKPS 309
 67 RTTSLYAEYNGAKOTVFC-IEPVSISSPTEVTHOY-----ORNPL 104

310 PGGPERLASM--AADTAHLHENVFTNGIYE-----TPDIKEKPMFI 353
 105 PSMSDKAKLUVSLWKEAGTDIDTMVAQKMIWEEVNGYKLHSIKRUGGASTDIK-----158

354 GMETGLPRNALGGYTA-----VAGVQ-AMVFSNSPTEVNLTEVDSGMTEAKDGGP 405
 159 SIEGKINK-AIEEYQKKPSFHNTVKTLGQSTLIDKNELNSEDPKVQNTA-----211

406 GPSFNRFYOFAGPHLAANPQTDRDGHLVLSQSFGSNTFSPVYLAICGFSAPILLRL 465
 212 ---NIDYRVTGNQLVLT-----NSNSKGSTLTKSAGTGP-VAYKK 251

466 FYLERCDAGAFTGGHDALKYVGTDFDSEIPCSCLCEKHTRPVCAHTTVHL-RQRMPRFQG 525
 252 AGLQTMAGALDKNTYA-KINBTGKS-LKTKKIDKESDIPETVHL-----DFGK 304

526 A-----TROPIGVFGTMSQYSDCPLGNYA-----PYLLRKPGDOTEAAKATM 570
 305 ALPSKDVTDDKDG-----SILDGIPIHSTKVTITKEKSVPDPMDITPMATIKAGETI 358

571 QDTYRATLRLFDIQLERBLDQ-----SVIUDHPT--FRRILDTRA 621
 359 SMTSKMRQKGQOLLEKIG-VETGTDLNNDNYSLAGNTKASDSPACEIVQEITTDKG 417

622 RIEQTTQFMKVLYVTRDYKIREGULSATHSMALTDFP-----YSGAFCPLTNFLVKT 675
 418 RA-E-TPEKIANALELGTYVTE-TKSSNGFNTFEKTVKELYKANQTVVALVTSNTYKQG 473

676 HLVYQDLAL-----SQCCHCVFYQQV-----EGNRFRNQFQPVLRFFVDL 717
 474 NQEITGETTLIKEDEXDTGNESQGAEFKAEXTIFTAKDQGQWKSEAFK-----TEL 526

718 FNGGETSTRSTVYLSE-GRVSAAPNPLGQ-----DAPAGRFDGDLARVSE-----764
 527 VKGTRASDETVTLALDERKNGQAVXRLAINEFWMDTETRPEGYTDEKPVSTKRDVNNE 586

765 --VDIRIVKRNRYV-----PSGNCTNLSEAA-----RARLVGLASAYQRQEKR 805
 587 KNAVITRDVTAKEQVTRGFDFKFAGSDGTAATGFNDLSFVSPLEGTXEITGAEDKA 646

806 VDMHLHGALGF-----LLKQFQHGLLPRGM---PPNSKS-----PNP 838

647 TTACNEQLGFDGYKGFENLPYGDVLEEEA--PEGFQKITPLIEIRSTFKENKDDYAKS 703
 839 QMWFTLQRNO-----MPADLUTHEETTAVKRTEYTAINFNP--PTCIGE 888
 704 EYVFITEGOKQKPMVTPYKLTNN-----FSVSNRMLYDLPKEKDSITS 754

889 LAQFYMANLILKCYCDESQYLTSLIITGARRPRDSSVLHWRKD--VTSAAIDIQ 945
 755 LATWAGKKNLNTLDPE-LVDRK-----RINLHEKEDWVVAANDVA- 799

946 AKALLEKTEUNPELWTATAFST-----HLVRAAMNQRMPTVLGISTSKYHGA 992
 800 TAAQAEQEDKARPVVIAETTALANEKETGKTMWLHKLTABQ-----VLDKSTIVLFNYV 853

993 ACNRYFQAGNWSGNGKVNQPLFEDTRPRELITACPREGFICPVTGSSGMBETLSD 1052
 854 YBNKVAEAGNE-----PVA-----KDASLNN 875

1053 QVRGIIVSGGANVQLATYATVRAVGR-----AQHMAFDWMLSLTDDFLARDL- 1102
 876 QAQ-----TNCUTIERHVSQTLKAHLDGSQTFTHGDMMDPDDQVTPKQFQ 923

1103 --FELHDQIQIQTLETPTVVEGALEAVILDE-----KTTAGD---GETPMLAFNED 1149

Db 924 GSKEAFETIYALLPDGTNKEI-WKSQKLEHEVYNDKEEFTKTVLAEKVDTGKYPEGTRKFTF- 982
 Qy 1150 SCBISHDTSNV-----LNTSGSNISGSTVPGLKREPPDE 1185
 Db 983 -TEINYEKDGNTNGKHEDELDKERSQTLTPEKEVPTIPSTPKQFQ 1024

RESULT 6

US-03-071-035-394 Sequence 394, Application US/09071035
 General Information:
 Applicant: Gil H. Choi
 Title of Invention: Enterococcus faecalis Polynucleotides and Polypeptides
 Number of Sequences: 496
 Correspondence Address:
 Address: Human Genome Sciences, Inc.
 Street: 9410 Key West Avenue
 City: Rockville
 Country: Maryland
 State: USA
 Zip: 20850
 Computer Readable Form:
 Medium Type: Diskette, 3.50 inch, 1.4Mb storage
 Computer: HP Vectra 486/33
 Operating System: MS DOS version 6.2
 Software: ASCII Text
 Current Application Data:
 Application Number: US/09/071, 035
 Filing Date:
 Classification:
 Prior Application Data:
 Application Number:
 Filing Date:
 Attorney/Agent Information:
 Name: A. Anders Brookes
 Registration Number: 36,373
 Reference/Document Number: PB369P2
 Telecommunication Information:
 Telephone: (301) 309-8504
 Telefax: (301) 309-8512
 Information for SEO ID NO. 394:
 Sequence Characteristics:
 Length: 1074 amino acids
 Type: amino acid
 Strandness: Single
 Topology: linear
 Molecule Type: protein
 US-09-071-035-394

Query Match Score 123; DB 4; Length 1074;
 Best Local Similarity 18.3%; Prod. No. 0.013; Pred. No. 0.013;
 Matches 206; Conservative 150; Mismatches 415; Indels 352; Gaps 52;

Db 105 PMSDIAKLVSVLWERACTIDTMVAOKMIVWEVNGYKLHSIKRGGASVDIK-----158

Qy 250 RVRNTAVARGAAHLAFADENHEGAVLPPDITYTFOSSSSGTTARGARRNDVNSTSCKPS 309
 Db 67 RTTSLYAEYNGAKOTVFC-1BPGVSIPTEVHRYG-----QRNPL 104

Qy 310 PSGGBFRRLSIM--AATDILHAEVFTGTYEE-----TPTDKEWMFII 353
 Db 105 PMSDIAKLVSVLWERACTIDTMVAOKMIVWEVNGYKLHSIKRGGASVDIK-----158

Qy 354 GMBGTDPRLNALGSYTA-----VAGVIG-AMVFSNPSALVTEVEDSGMTFEAKDGGP 405

Db 159 SIEBKINK-AIEEYQKKPSFHNTVKTILQGOSTTIDKNEELNSEFDKVQNTA-----211

Qy 406 GPSFNRFYOFAGPHLAANPQTDRDGHLVLSQSFGSNTFSPVYLAICGFSAPILLRL 465

Db 212 ---NIDYRVTGNQLVLT-----NSNSKGSTLTKSAGTGP-VATKK 251

Qy 466 FYLERCDAGAFTGGHDALKYVGTDFDSEIPCSLCEKHTPVCVAHTTVERLRQIMPRFQ 525

252 AGLTWAGALDKPNTAIKINVETKGS-LKIKKIDKESEGDIVPTEVPHL----D-FGK 304
 526 A-----TROPPIGVFGTMNSQYSDCDPLGNYA-----PYLILRKPGDQTEAKATM 570
 305 ALPKRDVTTDKDG-----SLDGIPHGKTVITTEKSVPDPYMDTTPTMAATIKAGETI 358
 571 QDPRATLERLFIDLEQRLDRGGAPCSSEGIS-----SVTVDPT--FRRIDLTLRA 621
 359 SMWSKNNMRQKGQILLEKTG-VETGTDWLNNTSAGNTFAIRKDSPAGEIVQETDTDERG 417
 622 RIEPTTQEMKVILVETDVKIREGLSEATHSMALTFDP-----VSGAFCPTNLVLRVT 675
 416 RAE-TPEELANALELGYYVTE-TISSNGPVNTKPKTVEKLYANQTYALVTSVNVKGQ 473
 676 HLAVVQDIAL-----SCQCHCVFYGQQV-----BGRNFRNQFQPVLRRFVDFI 717
 474 NQEITGETTLTKEKDGTGNESQCKAEFGKAETYLFTAKDQGAVKNSEAFK-----TEL 526
 718 FNGGFISRSITVTLSE-GPVSAPNPYIGQ-----DAPASRTFDDLARVSVE---- 764
 527 VKGTXASDVTLLADERNQAVKHLAINEWYQETKAPEGYTLDTKYPISVXKVVDNNE 586
 765 -----VDIRTRVNRRVV-----FSGNCNTNLSEAA-----RARIYGLASAYQROEKR 805
 587 KNAVITRDVTAKEQVIRGFDFKKFKAGDGAETGFDNLSSFKVSPLEGXTITEAEDKA 646
 806 VDMFHGLGE-----LLKOFHGLLPPRGM-----PNPSKS-----PnP 838
 647 TTACNEOLQGFDGYGKFKENLPYGDYLLEBIA---PEGFQKITPLTEIRSTFKENKDDYAKS 703
 839 QWFYFLQLQNO-----MPADKLTHBETTIAAVRKFTEYYAATFINAP--TCIGE 888
 704 EYVFITTEGQKQDVKONTVPIVKLTNE-----FSVSLNRMLMYDPEKEDSLTS 754
 889 LAQFYMANLILKCYCQDHQSQYLINTLTSLTGARRPRDPSSVLLHWRIRD---VTSAADETQ 945
 755 LATWDGNGNKLNTDFT-LVDKL-----RYNDHEIKEWVYVAQADVEA- 799
 946 AKALLEKTEENLPMLWNTAFTST-----HLVRAAMNQRPMVVLGISISKYHGA 992
 800 TKAQEKEDKAPVIAETTATLANKEKTKTGKWLHKLTAEQ-----VLDKSLTFLNTY 853
 993 AGNIVFQAGINWSGLNGRNVCPLFTDRTRFLIAQPRGGFICPTGPSSGNRNTLSD 1052
 854 YENKYAEBAGNE-----PTA-----KZASLNN 875
 1053 QVRGIIVSGGAMVQLAIAVTVRARG-----AQHMAFDWLSLTDDEFLARDL-- 1102
 876 QAQ-----TVNTCTERVISQTKAHBEDGSJSTFHGDNDMFDDVSVTHDVLD 923
 1103 -BEHDLQIOTLTELPTWVGEALKILLE-----KTAGD-----GETPTNIAFNFD 1149
 924 GSKEAFETIYLALPDGTNEIWSKGKEHEVYDKEFTKTKVIAEKVDTGKYPEGKPKTF- 982
 1150 SCEPHADHTSNV-----LNISSNSNISGSTVTPQKRPEDDE 1185
 983 -TEINYEKDGVNVGKHNEDIKEKSQLTIPKVEVPTSPPKQE 1024

103 SAAAPNLTRACNAARERFGFSRQGPPVUDGAVETTGAEICTRGLEPENTILLYVVTALF 162
 207 ESP-----VTEBFG-----
 Db 163 KEMAVMCNVPLHYGGDIVHINHGDVIRIPLFPVQLFMDYNTLVPDPENTHRSIGEGF 222
 232 -----NGFVREG-----GVVVLKPL-----SRALDGDRVH-----GV 261.
 223 VFTPTFINTGLCLHINDCIAPIAVARLVRNTAVARGAHLAFDENHEAVLPDITYT 282
 262 IRRASVNNDG-----ATPGUTVPSAQAQKVLRAYXQALDPS-AVQ 303
 Qy 283 YFOSSSSCTTAGRNRDNVTSKSPSGGBERRLASIMADTAHLAEVLFNTGYEET 342
 Db 304 YVBLHGT-----PGDPLLEAAAGVLSARRADEP-----LVGSA 342
 Qy 343 PTDIKWPMFIGMEGTPLRNLAGSTARVAGTIGAMVEPSNALSITYVEDSGNTCEAKD 402
 Db 343 KRNVGHLSGAAGVGVLKLTLAG-----RRRIPASLNTPTHPDPL-----DTLGLDVD 394
 Qy 403 GGPGPSEFRYQFAGPHLAANPOTRD-----CHVLSSTOSSTSSNTFSDV 448
 Db 395 G-----LREWPPDPRELLAGYVSFGGMGTVAHVVLSEGPAQGGSGPGID 438
 Qy 449 YLALICGFQAPLARLFLYERCDAGA-----FTGGHGDALK-----YTGTDFSEI-P 496
 Db 439 EETPV-----DSGALPFTVTTGRGEAQLRAQRLHEAVEADELAP 480
 Qy 497 CSLCEK-HTRPYCAHTV-----HRLQRMRVERGQATROPIGVGTMSOYSDCDLG 548
 Db 481 AALLRSVLTTRTFTHRVSVLADRARIIDGLGALAAGTPAAGVVTGT-----528
 Qy 549 NYAPYLIRKGQDEAAKATMDTYRATLERFIDLEQERLLDRGAPCSSEGSSLSSVIVD 603
 Db 529 -----PAPG-----RLAVLP-----SQQGQAORTMGCMELYAA 555
 Qy 609 HPTFRRILDTLRARIEQTTOFMKVLYETRD-----YKREGLSSEATHSMALT 656
 Db 556 HPAATAFPAVAEELDPLDRPAELVAAGTIDUDRVTHTQPALFAEVHLRVBSWGT 615
 Qy 657 FDPYSGAFCPITNLFVKETHLAVYDOLALSOCHCVCYGOOEGRNFENQFQFVLRERFD 716
 Db 616 PDLAGH-----SYGEISAAHVAGY-----LSLRDA-----ARLYAARGRMQALP-----656
 Qy 717 LFNGCFISTRSIYTLLSECPVSPAPNPTQDADAGRITDRLARYSEVIRDIVRKNRY 776
 Db 657 -EGGAM-----VVAEASEEVL-----PHU-----AGRENELSLAVERN-----694
 Qy 777 FSGNCNTNLSEAAARLVLGLASAYORQEKRVDMMRGALGFLKOFHGLF-----825
 Db 695 LAG-----AERAVLDVAELREGRTTRKLSSVHA-----FHSPLMPEMLDDPRVV 741
 Qy 826 -----PR-----GMPPNSKSXPNPW-----FWTLLQRNQMPADKLTHBTITIAAV 866
 Db 742 EELDEQEPYDVSVTGTGLPVTIA-----GWTDDBYW-----VTDQV-RRPVRFDV 787
 Qy 867 KRFTEYAAINFNIIPP-----TCIGELAOPMANILKYCDSOULINTLTSITGARPRD 924
 Db 788 -RTLEESGATDFLEGPDGVCSSAAAADSY-----RDQRAATAVSALRKG-RP-E 833
 Qy 925 PSVHLWIRKDVTSAADIEQAKALLEKENLPELWTAFTSHLYRAAMNQRPMTVLLGI 984
 Db 834 PQSLLAALTIVFVRGHDVD-----WTAHGSTGTGVPL-----PTTAFQR 874
 Qy 985 SISKYHGAAGNRVRQAGNWGLANGKNCVPLTFDRTRPFIACPROGFCIPVGPSSG 1044
 Db 875 ERHWDGQARTAAPLTAGR-SGTGAG-----TGAAAG 905
 Qy 1045 -----NRFTTLSDQVRGJIVVSGAMVQLAIVATVRA-YCARA 1081
 Db 906 VTSGEGEGEGEAGAGGGDRPARHETT-----ERVIAHVA-----VLEYDDPITVELGUTP 957
 Qy 1082 QNAFDWLS-----LTDDPBLARDLEELLDQIQTLETPTWVGALEAKVILLDEKITA 1135
 Db 958 KEGFDSUMSVLRLNVAUDDCLRLPSGLJDFH-----PT-----RALAA-HJSDLITG 1006
 Qy 1136 GDGETPTINLAFWDFDSCPS-----EDTTSANVLNSIGNSNISGSTMSPGLKRPDEDELFDL---1189
 Db 1007 GSGETG-----SADGTPATPADTTAEPIALIG--MACRYPGGTSPED-LWFLVAEG 1056
 Qy 1190 -----SGIPIKHG 1197
 Db 1057 RDAVSGLPTDRG 1068
 Qy RESTLT 9
 Db US-07-612-734-C-4
 Qy Sequence 4-A Application US/07642734C
 Db Patent No. 582513
 Qy GENERAL INFORMATION:
 Db APPLICANT: Katz, L
 Db APPLICANT: Donadio, S
 Db APPLICANT: McAlpine, J B
 Qy TITLE OF INVENTION: Recombinant DNA Method for Producing
 Db NUMBER OF SEQUENCES: 27
 Qy CORRESPONDENCE ADDRESS:
 Db ADDRESSEE: Edward H. Gorman
 Db STREET: Abbott Laboratories D377/AP6D-2 One Abbott
 Db CITY: Park Rd
 Db STATE: Abbott Park
 Db COUNTRY: US
 Db ZIP: 60064-3500
 Qy COMPUTER READABLE FORM:
 Db MEDIUM TYPE: Floppy disk
 Db COMPUTER: IBM PC compatible
 Db OPERATING SYSTEM: PC-DOS/MS-DOS
 Db SOFTWARE: Patent In Release #1.0, Version #1.25
 Qy CURRENT APPLICATION DATA:
 Db APPLICATION NUMBER: US/07/642,734C
 Db FILING DATE: 17-JAN-91
 Qy CLASSIFICATION: 435
 Qy ATTORNEY/AGENT INFORMATION:
 Db NAME: Dancikers, Andreas M
 Db REGISTRATION NUMBER: 32552
 Db REFERENCE/DOCKET NUMBER: 4952-US-01
 Qy TELECOMMUNICATION INFORMATION:
 Db TELEPHONE: 708-937-9396
 Db TELEFAX: 708-938-2623
 Qy INFORMATION FOR SEO ID NO: 4:
 Db SEQUENCE CHARACTERISTICS:
 Db LENGTH: 3567 amino acids
 Db TYPE: amino acid
 Db TOPOLOGY: linear
 Db MOLECULE TYPE: protein
 Db US/07-642,734-C-4
 Qy Query Match Score 106.5; DB 2; Length 357;
 Db Matches 23; Conservativeness 104; Mismatches 384; Gaps 62;
 Qy Qy 125 CGPDPVDAVETTSAE-----ICTR-----LGEPEANTILVVTALKEAVPMCNVFLH 174
 Db 1165 CGPDDLLAAVEAGASAVICQDAAALEEALGDBP-----VTAI-----VH 1205
 Qy 175 YGGDIVHNGDVTRIPFVYOLFMDPVNRLYPDFPATHTHSIGEGRVPPFXTNGGLC 234
 Db 1206 AGTL-----TNPG-----SISKEAPEEF-----ATTAIAKPAIALL 1234
 Qy 235 FLIHDCVIAVMAYALRV-----RNVTAVARGRAHL-AFDENHEGATLPPDITYT 282
 Db 1235 AVL DEVGRAREVYCSVAGWGGAGMAYAAGAYLALAEIHR-----1282
 Qy 283 YFOSSSSGGTTARGARRNDWSNSKPSBG-----GF--ERRLASIMAAADTLHA-EVFIN 335
 Db

1283 --- -ARGRCSVTAWTIPWALPGEAVDDCYLRLRS-TSADRAMEWTWERTVIA 1330
 336 TGIYIETPTDKEWPF1-MEGGDPRLNALGS-TARVGVIGANVFSNSALYLTEVED 394
 1331 AGPVSVAVADY-DWPVLSSEGFAAERP-----TAFAELAGR 1365
 395 SGMTAEKDGGECPSPNFRYQFAG---PHLAANPQTDDGHYLSQSSTGSNT 444
 1366 GSQAREPDPS-GTPEPAQDLAGLSPDEQENILLELVANAVA--WLGHEBDAEINR 1420
 445 PSVDYLALICFGGAPPLARLFYLERCDAGAFTGGHGDALKYVGTDFEDBIPCCLCBKH 504
 1421 RAFFEGLDSLNAMAIKRKL-----SASTG-----LRLPASL-----1452
 505 RPPVCAHTTVHLRQ-RMPRFQATRQPIQYFGTMNSQYSDCBL-----GNYA 551
 1453 --VFDFHPTVTAALOHLARLVLGDAQAAVRVGAAAD---ESEPIAIVGIGCRFPGGIGS 1506
 552 PYLIR-----KPGDQEAEAA-----566
 1507 PEQLRVEVLAEGANLTGFPADRGNDIGRLYHPDPDNPGTSVYDGGFLTDAAFDPGFFG 1566
 567 --- -KATMQDTYRATLERFLIDQBLGAPCSSEGGLSSVITDHPPTFRR 614
 1567 ITPREALAMDPOQRMLMLETAEVERAGIDPDAARGTDTGTVFGVMNGQSYMOLLAGEER 1626
 615 I-----LDTLPARIET-----TQEMKVILVETDRDYKIREGLSEATHSMAL 655
 1627 VDGQGLGNSASVLSRIAYTFGGWGPALTVDIACSSLVGI-HLAMOALRIGECSDL 1684
 656 T----FDPYSGAFCPITPLVKRTHLAVQDIALSOCHCVFYQVEGRNFNN-----704
 1685 AGGTVTYNSDPT----FVDFSTOR-----GIA-SDRCKKASARAQGFALSEGYAAL 1731
 705 QFQPVLRRFVDFLNGGFISTRSTIVTLLSGPVSAPNPQLQDAPAG---RTFDGDLAR 760
 1732 VLEPLSRAR---ANG---HOVLAVLRLGSANVNDGASNLGAENGPSQERYRQALAA 1782
 761 VSVEVIRDIRYKRNRYVFSGNCTNLSEAAARLVLGASAYORQER-----VDMHLH- 810
 1783 SGVPA-ADVDV--VEAHGIGTLELDPIEG---ALIATYQDGRPRPLRQLGSVKTIGHTQ 1836
 811 ---GALGFLIKQOPFHGLIIFPRGMPNSKSPNPFNWTLQRMQPAKLTHEITIA 864
 1837 AAGAGAVKIVKUAMHGMLPRSHADELSPHDW-----ESGAVFVLREVPVPA 1887
 865 AVKRFTEEYAAINFNLPPTCIGLAFQYMANLILKYCDHSQYLINTLTSITGARRPD 924
 1888 GER---PFRAGSVSFSVSCGT-----NAHVTEVEAPEQ-----EARTERG 1925
 925 PSSVLHWRDVSAADETQAKALLEKTNLPEL-----WTRAFTSTHL-VGAANNQR 977
 1926 PLPFVLSGRSEAVAA---QARLAEEHLRDTPELGLTDAAWTLATGRARFDYAAV-- 1978
 978 PMVVLGIGISISKHYGAAGNNRVFQAGNWSGLNGKNCVPLFEDTRRFTIACPRGGFICP 1037
 1979 --- -LGDDRAGVCABDAAEGR-----PSADAVAP 2005
 1038 VTGPGSSGNRETTLSDQYRGIIVSGGMVQLAIVATYVRAVGRARQHMAFDDWLSLTDEF 1097
 2006 VT---SAPRKPYL-----VFGOGAQ-----WVG-----2026
 1098 LARDLLELHQDIIQTL-----ETPMVTEGALBAVKILDEKTTAGDG 1138
 2027 MARDLLESSEVEAEMSRCAEALSPTDWW-----KULD-VYRGDG 2055
 RESULT 10
 US-08-439-009A-4
 ; Sequence 4, Application US/08439009A
 ; Patient No. 6004787
 ; GENERAL INFORMATION:

APPLICANT: Donadio, S
 APPLICANT: Katz, L
 APPLICANT: McAlpine, J B
 TITLE OF INVENTION: Method of Directing Biosynthesis of Specific Polyketides
 NUMBER OF SEQUENCES: 27
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Steven F. Weinstock
 STREET: Abbott Laboratories D377/AP6D-2 One Abbott Park Rd
 CITY: Abbott Park
 STATE: IL
 ZIP: 60064-3500
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-PC/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/439,009A
 FILING DATE: 11-MAY-1995
 CLASSIFICATION: 435
 ATTORNEY INFORMATION:
 NAME: Casuto, Diane
 REGISTRATION NUMBER: 40,943
 REFILE DOCKET NUMBER: 4952.US.D1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 847-938-3137
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3567 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-439-009A-4
 Query Match Score 106.5; DB 3; Length 3567;
 Best Local Similarity 19.9%; Pred. No. 6.2;
 Matches 237; Conservative 10; Mismatches 384; Indels 465; Gaps 62;
 125 CGGPPVDAVETTGAE-ICTR-----LIGLEPENTILLYVNTALFKEAVFMCNVFLH 174
 Db 1165 CPGDDLLAAVEBAGASAVVCAQDAALREALGDEP-----VTA-----VH 1205
 175 YCGLDIVHNGDVIRIPLFLVQLMPDVNRVLPDPFNTHRSYEGFVYPPTPYNTGLC 234
 Db 1206 ATGL---TNG-----SISEVAPEEF-----AETIAAKTALL 1234
 235 HLIJHDCTVAPNAVALRV-----RNVTAYVARGAHL-AFDENHREGAVLPDPDITYT 282
 Db 1235 AVL-DEVLGDAVEREVYCSCSAGVWGGAGMAYTAGSAYLDALAEHHR-----1282
 283 YFQSSSSGTTARGARRNDVNTSKPSPSG---GF---ERRLASTIMAADTAHLA-EVIFN 335
 Db 1283 -----ARGRECTS/VAWPALPQGAVDDGylRLGRLS-ISADBMRTWVERLA 1330
 336 TGUYETPTDKIWKMPFI-GMEGLTPRLNALSYTARVAVGIVGMVFSNALSALTEVED 394
 Db 1331 AGFVSVAVADV-DWPVLSSEGFAAERP-----TALFAELAGR 1365
 395 SGMTBEAKDGGPPSPNFRYQDAG-----PHLAANPQTDRDGHVLSQSTGSSNT 444
 Db 1366 GGQAEBAPDS-GPTGEPQAQLAGLSDEQQENLLELVGHESSAEINV 1420
 445 FSVDYLAICGGAPILARLFLYLERCDAGAFTGSHGDALKYVGTDFSEIPCSLCERHT 504
 Db 1421 RAFFSELGLDSNAMALKRKL-----SASTG-----LRLPASL-----1452
 505 RPVCAHTTVHLRQ-RMPRFQOATRQPIGVFGTMNSOISDCDPL-----GNYA 551
 Db 1453 -VFDPHPVTLAQLHRLAVGADAOAVRVTGAAV-----ESBPIAVGIGCRFFGGIGS 1506

RESULT 11
US-07-731-157A-7
Sequence 7, Application US/07731157A
; GENERAL INFORMATION:
; Patent No. 5457032
; APPLICANT: Quax, Wilhelmus J.
; ADDRESS: Cooley Godward Castro Huddleson & Tatum
; STREET: Five Palo Alto Square, 4th Floor
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; DB 552 PYLILR-----KPGDQTEAA-----
; DB 1507 PEQIWRVLAEGANLTGCPADRMWIDGBLYHPPDPNGTSYVDKGELTDADDPDGFFG 1556
; QY 567 -----KATMDQTYRATERLRLFDLEDERLDRGAPOSSEGLSSVIVDHPTEERR 614
; DB 1567 ITPREALANDPQQFLMLMATEAWERAGIDPDLRGDTGTVGMNQSYMLLAGBAER 1626
; QY 615 I-----LDTURARIEQT-----TDFMKVLVEFRDYKREGLSATEATHSMAL 655
; DB 1627 VDGYQGLGNASVLSGRSLAYTPEGWGPALTVIDACSSLVGT-----HLMQALRRGBCSLAL 1684
; QY 656 T-----FDPYSAFCPTINFLVYRTHLAVQDALSOCCHCVYGOOEGRNFRN 704
; DB 1685 AGGYTVMSSPYT-----FVDSFQR-----GLA-SDGRCKAFSARADGFALSEVAAL 1731
; QY 705 QFPVYLRRRFVDSLNGGF1STRSITVTLSGPVSAPNFTLQPAAG-----RTFDGDLAR 760
; DB 1732 VLEPLSRAR----ANG----HOVLAVLRLGSANQDGASNLGAAPNGPSQERVIRQALAA 1782
; QY 761 VSVEVTRDVKNEVVSGNCNTNLSEARARLVLGASAYQREKR-----VDMHLH- 810
; DB 1783 SGVP-ADV-----VEAHOTGTGEQDPEAG-----ALIATYQDDDRPURLGSKTNTIGHTQ 1836
; QY 811 -----GALGFLKQFHGLLFPRGMPNPKSPNPWFTLQLQNONPADKLTHEETTIA 864
; DB 1837 AAGAGAVKVVLAMRHGNL-PRSLHADELSPHDW-----ESGAEVTRREEWPWA 1887
; QY 865 AVKRTEETTAANFINLNPPTCIGBLAQFYMANILKYCDHSQYLINTLTSITGARRPD 924
; DB 1888 GER---PRAGVSSFGVSCST-----NAHTIVEAPEQ-----BAARTERG 1925
; QY 925 PSSVHLWIKRDVTSADIEQTAQAKALLETNLPEL-----WTIAFTSTHLL-VRAAMNQR 977
; DB 1926 PLPFVLSGRSEAVVAA-----QARLAELHLDTPLGLTDAAWTLATGRARFDVRAAV--- 1978
; QY 978 PMVYLGISISKYHGAGNNRVFQAGNWSGLGKRNVCPLFTPRTTRFIACPRGGFTCP 1037
; DB 1979 -----LGDDAGCAELDAEGR-----PSADAVAP 2005
; QY 1038 VTGPSSGNRETTSQVQRGTTIVSGAMVQLAIATYVRAGBAAQHWAFFDDMNLSLTDEF 1097
; DB 2006 VT---SAPRKPV-----VFPQGAQ-----WVG----- 2026
; QY 1098 LARDLEELHQIQT-----EPPTVGEALEAKILDKXTAGDG 1138
; DB 2027 MARDLESSEVFVAESEMSMRCAEALSHTDW-----KLLD--VVRGDG 2065

US-07-731-157A-7
Sequence 7, Application US/07731157A
; GENERAL INFORMATION:
; REFERENCE NUMBER: US/07731157A
; FILING DATE: 1990-10-09
; PRIORITY NUMBER: EP 902009562
; FILING DATE: 1990-04-18
; ATTORNEY INFORMATION:
; NAME: RAE-VENTER PH.D., BARBARA
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: G BRO-027/7/0015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-494-7622
; TELEFAX: 415-857-0653
; TELEX: 380816 COOLEY PA
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 774 am.no acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Pseudomonas species
; STRAIN: S833
; US-07-731-157A-7

Query Match 1.7% ; Score: 106 ; DB: 1 ; Length: 774 ;
Best Local Similarity 20.2% ; Prod. No.: 0.4 ;
Matches 145 ; Conservative: 79 ; Mismatches: 251 ; Indels: 242 ; Gaps: 37;
Qy 258 ARGIAHALPENHNGAV-LPDITTYTQSSSCTTARGARENNDVNSTSKSPSGGER 316
Db 183 ANALKLRLTDGGDLICPGVEARLEADLALRPAVDALLKAMGDDASDAAAGGSNN 242
Qy 317 -----RLAS---IMAADTAHLAEVIFTGTYPEBTPDIKEWMFINGETGLRLNALGS 367
Db 243 WAVAQRTAAGRPLAGDHRVFEI---PGMYQHHLACDRDFM-IGL-TVP--- 289
Qy 368 YTARYAVGIGAMVSPNSA-----LYLTEVEDSGSMTE----- 399
Db 290 -----GVPRPFPAHNGKVAYCATHAFMDIDHLYLEFAEDERTARGNEFEPVAMRD 343
Qy 400 -AKDGGPGSENFYQFQGPHLANKPQTDRGHVLISSQSTGSSNTTEFSVDYLAICGFG 457
Db 344 RIAVRGGADEFDIVETRGPVLAGDP---LEGALHTLRSVQAEATDLSFDCHTRMP-G 398
Qy 458 APILLRLFLERCDAGATG--SH---GDALKYVITGDFDSEPCSLCDEKTRPVCAHT 511
Db 399 ASTVQLY-----DATRGWGLDHNLVAGDVAGSIGIYRVARVPSRPRENGNLPVPGKS 452
Qy 512 TVHRL-----QMRPFGQOTROPICGVFTGMNSQY-----SDCDPLGNAYP----- 553
Db 453 EBEHWGWHIPHEANPR-----VDPDPPGLITANNVVAHDHPDYLCTDCHP-----PYRAE 504
Qy 554 -----LILRKQGDOEAAKATMQT-----YRATLERLFI--DLEQERI----- 590
Db 505 RIMERIVASAFAVDDAAAHADTLSPHYGLRLEAIGQGSLPAEEERQTLIAWDGR 564
Qy 591 LDGPGCOSGSSGSSVTDHDTLARISQTTQFMKVLYTREDYKIREGSEA- 649
DB 565 MDAGSQAAA-----XAFRAL-----TRVTTARSLEQAT 596
QY 650 THSMALT-----EDPYGARGCITINFLYKTRHLYQ-----DIALSOCHCYFQQVEGRNF 702
DB 597 AHFAVAPPGVSPQGQMVNPYPT-LRNNDAGMKGWSNDEALSEPSV-ATONUTGRW 654
QY 703 RNQFQPVLRREFDVLNGGKJISTRIVTSEGVPYS-----NPTLQDAP 749
DB 655 GEEHPR-----RPTPLSAQPWAALL---NPVSPRPIGDGTIVLANGLVPSAGEPAT 705
QY 750 AGRTEGCDLARYSVETDRVKNVUVF-GNCTNLSEARARLYGLASAYCROKRVDM 808
DB 706 YG-----ALSRYFDDVGNWDN-----SRVWV 726

Y 809 LHGALGFLLKQFHGILFPGRMPPSKSPNPKW----FWTLLQRNQMPADKLTHEEI 860
 Y 727 FHGASG-----HPASPHYADONAPWSICAMVPMLYSWDRIAAEAVTSQEL 771

RESULT 12
 Sequence 7, Application US/08541780

GENERAL INFORMATION:
 Patent No. 5035531
 APPLICANT: Obax, Wilhelmus J.
 APPLICANT: Misset, Onno
 APPLICANT: Van der Laan, Jan M.
 APPLICANT: Ienting, Herman B.M.
 TITLE OF INVENTION: Mutated beta-lactam acylase genes
 NUMBER OF SEQUENCES: 50
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: COOLEY, GODWARD CASTRO HUDDLESON & TATUM
 STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR
 STATE: CALIFORNIA
 COUNTRY: USA
 ZIP: 94306

COMPUTER READEABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/541,780
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/731,157
 FILING DATE:
 APPLICATION NUMBER: EP 90200962
 FILING DATE: 18-APR-1990

ATTORNEY/AGENT INFORMATION:
 NAME: RAE-VENTER PH.D., BARBARA
 REGISTRATION NUMBER: 32,750
 REFERENCE/DOCKET NUMBER: GBRO-027/00US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-494-7622
 TELEFAX: 415-857-0663
 TELEX: 380816 COOLEY PA

INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 774 amino acids
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOOTHEICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Pseudomonas species
 STRAIN: SE83

IS - 08-541-780-7

Y 400 -AKDGGPGPSFNRFYQAGPHLAANPQTDRDGHVLSQSTGSNTTEFSVDYLALICGFG 457
 Y 344 RIAVGGADREFDIFTTRGPVAGDP---LEGALTIRSVPATEIDSFDTTRMP-G 398
 Db 458 PILLARLFLYLERCDAGATG-GH---GDALKYVTGTDFDSIPICSLCEKHTRPYCAHT 511
 Db 399 ASTVAQLY----DATRGWGLIDHNLYAGDVASIGHLVRVPSPRENGNLPVPGWS 452
 Y 512 TWRHL---.QRMDFREGATRQPIGVFTMNSQY----SDEDPLQNYAPY--- 553
 Qy 453 GEHENRGWLPEAMEV---VDPGGLIVTANRVRVADHPYLCTDCHP----PYRAE 504
 Qy 554 ---LILRKIGDQEAAKATMDT----YATLTLIFI---DLBQERL---- 590
 Db 505 RIMELVVASPAFAVDAAHADLSPHYGLARPLAELGIQSLPLPEELRQLLIANDGR 564
 Qy 591 LDRCAPCSESEGGLSSVIVDHPTFRRLDTLRARIEQTQTQFMKVLYVETRDYKIREGLSEA- 649
 Db 565 MDAGSQASA----VNAFRAL---- 596
 Qy 650 THSYSLT---FDPYSGAFFPTINLVKRTHLAVYQ---DLALSQCHCVYQGOVEGRNF 702
 Db 597 AHPPFAVPPGVSPQQVMWAVPT-LLRNDAHNLKGNSWDEALSEASV-A7ONLTGRW 654
 Qy 703 RNQFQPVLRBRPFVLDENGPISTRSITVLSGPVSAP----NFTLGQDAP 749
 Db 655 GEEHHP---RFTHPLSAQFPANWALL----NPVSRPTGGDGTVLANGLVSAGPAP 705
 Qy 750 AGRTFDGDLARVSVEVDIRKARVVPS-GNCNTNLSEARARLVLGLASAYQRQERKVDM 808
 Db 706 YG----ALSRYVPDFGVNDN----SRWV 726
 Qy 809 LRGALGFLLKOFHGILFPGRMPPNSKSPNPKW----FWTLLQRNQMPADKLTHEEI 860
 Db 727 FHGASG-----HPASPHYADONAPWSICAMVPMLYSWDRIAAEAVTSQEL 771

RESULT 13
 US-09-107-532A-3855
 Sequence 3855, Application US/09107332A
 Patent No. 6583275
 GENERAL INFORMATION:
 APPLICANT: Lynn A Doucette-Stamm and David Bush
 TITLE OF INVENTION: NUCLEAR ACID SEQUENCES RELATING TO
 ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 NUMBER OF SEQUENCES: 7310
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 STREET: 100 Beaver Street
 CITY: Waltham
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02354
 COMPUTER READABLE FORM:
 COMPUTER: PC
 COMPUTER TYPE: CD/ROM ISO9660
 OPERATING SYSTEM: <Unknown>
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,532A
 FILING DATE: 30-Jun-1998
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 APPLICATION NUMBER: 60/051571
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Arinello, Pamela Deneke
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781) 893-5007

INFORMATION FOR SEQ ID NO: 3855:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1095 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: YES
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 NAME/KEY: misc feature
 LOCATION: (B) LOCATION 1...1095
 SEQUENCE DESCRIPTION: SEQ ID NO: 3855:

Query Match Score 105.5; DB 4; Length 1095;
 Best Local Similarity 18.9%; Pred. No. 0.86;
 Matches 145; Conservative 112; Mismatches 298; Indels 211; Gaps 35;

551 APYLURKPGDQEAAKATMDTYRATLERLFLDEQERLDPGCSSEGLS-----S 604
 360 APTIDTPMFTTIAKABETIYTSKAQEKQIILDKSG-VETGSLLWNNSYAGNTA 418
 605 VIVDHPPTERRILD-TLRARIBOTTQFMKVLYKIREGLSEATHSMALTFDP--- 659
 419 IRKDSPTEGEIVQMTTDENGHAETPKIANALEGTYYVTE--TRASHGFVNNTPKPVKE 476
 660 - YSGACFPIITFLVRTHLAVYQDAL-----SQCHCVFYQGQV-----EGR- 700
 477 LRYANOFVVALVTNSVKCONCQEYGETLTKDGDKAQGKAVVEGTEYTFLFTAKDGKA 536
 701 -NFRNQPDPVRRFDLFGSFIRSITVTLSEGVPVAS----NPLTGQD--APAGR 752
 537 VKSEAKP-----EMVKGTKAASDETVTLALDENQAAVKHLAINETWQEKIAPEGY 589
 753 TPDGDLARVSVE-----VIRDIRVKNRV-----FSGNCNTNLSEAA----- 788
 590 TDETCKPVSKVDDNEKNAVTRDYTAKEQIIRFGDFPFKPGASAAGTAETGNDLTf 649
 789 -RARLVGLASAYQROERVDMLHGAIGF-----LKKOFGHGLFPGM----- 829
 650 KVSPLEBGNEITGQAEADATTAVNEEQGQFPLVYGGYLEVEA--PBFQKTT 706
 830 PPNISKSP-----NPQWFNTLQRNQ-----MPADKLTHEEITIAAVKRFTEE 872
 707 PLIEIRSTFKENKEDEVSEYVPTTEDQKQPIKTVTPYKLTN-----KAFSVS 757
 873 YAIINFNL-----PTCIGELAQPYMANILKVKDSQVLINTLTSITGARRPRDSSVLU 930
 758 LNRLMLYDLPPEEDSLTSLATKDGKELTSIDSTE-LVDKL-----SYNLH 803
 931 WIRKD---VTSADLIEQAKALLEKTENLPELWTTAFTSTLVLRAAMN-----OPRMV- 980
 804 EKEDWVVAQALDVDA-TKLAQEKDEKAKEP-VIILETSATLANKEKTGWIKQHLTAE 861
 981 -VIGISISKYKGAGNNRVFOAGNWGLNGKKNVCPPLFTFDRTRRFLIACPRGFCICPT 1019
 862 QVNRTKTVLNTYENKEAEGD-----KPA 889
 1040 GPSSGNRETTLSQVRGTIVSGGAMYQLAIYATVVRAGFAEAQHMAFDWNLSLTDDEFLA 1099
 890 -----KDVSLNNQAQTVCSTVHHVIQTKAHLENG----SQTTHGDVYDMFDVDSIT 939
 1100 RDL---FELHQDIIQLETPWVEGALEAKVILDE-----KTF---AGDGTPTNL 1144
 940 HDVLGSKSEAFTILYALLPDGTKNEWKSGKIEYEYNDKRFKTVLAKVDTGYPEGT 999
 1145 AFNDSCPSPSHDTSN----VNIUNGSNISGTSVPLKRPEDDE 1185
 1000 KFFFAEINYDKDGTINGRNHEDLKRSQTLTPEVNTPLSPKQE 1045

RESULT 14
 US-08-633-760-46
 ; Sequence 46, Application US/08633760
 ; Patent No. 5,804,429
 ; GENERAL INFORMATION:
 ; APPLICANT: NIWA, MINEO
 ; APPLICANT: SATO, YOSHIMASA
 ; APPLICANT: FUJIMURA, TAKAO
 ; APPLICANT: ISHII, YOSHINORI
 ; APPLICANT: NOGUCHI, YUJI
 ; NEW CEPHALOSPORIN C ACYLASE
 ; TITLE OF INVENTION: A NEW CEPHALOSPORIN C ACYLASE
 ; NUMBER OF SEQUENCES: 64
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAYER & NEUSTADT,
 ; ADDRESS: 1755 JEFFERSON DAVIS HIGHWAY, SUITE 400
 ; CITY: ARLINGTON
 ; STATE: VIRGINIA
 ; COUNTRY: USA
 ; ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/633,760
 FILING DATE: 01-MAY-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: OBLON, NORMAN F.
 REGISTRATION NUMBER: 24,618
 REFILE NUMBER: 18-929-0 PCT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 413-3000
 TELEX: (703) 413-2220
 TELEX: 248855 OPAT UR
 INFORMATION FOR SEQ ID NO: 46:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 774 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-633-760-46

Query Match Score 105; DB 1; Length 774;
 Best Local Similarity 20.7%; Pred. No. 0.5;
 Matches 146; Conservative 76; Mismatches 241; Indels 242; Gaps 37;
 QY 109 LTR--ACNAARERFGFSRCQGPVDAVETGAE-ICTR--IGLEPENTILYLYVTAL 161
 Db 70 LTRKAGLGRAAWLG--AEAAEADLIVRRLGMEXYRRDDEALGVEAKD-----N 117
 QY 162 FKEFRAVFCNVFLHYGGGLDIVHINHGDVTRPLFPYQOLMPDNRYLPDPFNTHH----R 216
 Db 118 LRAYVAGYNAAASGA-----EL-PVEYGLQAE--PEPWDPWHSIAVMR 159
 Db 217 SIGEGFVFYPTPYNTGFLCHLHDTCVIAPMAYALRVNTAV---ARGAAHLAFDENHEG 272
 Db 160 RLG-----LIMGATWFKWIRMLAIPVVGANALRLRYDDGGRD 197
 QY 273 AV-LPPDITYTYFQSSSSGTTARGARRDNSTSCKPSGSGFER----RLAS--IM 322
 Db 198 LLCIPPGAAEADLTLRPAVDALLKANGDASDAAGGSNNWAVAPRTATGRPIL 257
 QY 323 AADTALBAVIFNTGAYETPTDIREWPMFQMGETGLPRINALGVSYAVGIGAMVFS 382
 Db 258 AGDPHRVFEI---PGMYAQMHLACDRDM-1GL-TVTP-----GVPGFPHFA 298
 QY 383 PNSA-----LYL-TEVEDSGMT-----AKDGGPGPSFNRF 412
 Db 299 HNGKVAICVTHAFMDIHDLYEQPAGEGRTARFGNDFEPVAWSRDRIAVRGGADREFDIV 358

413 YQAGPHIAANPQTDRGHVLSSQSTGSSNTFSYDYLALICGFAQILLARYLERCD 472
 359 ETRHGPVIAGD--RDGAALTLSQFAETDLSPDCLTAMP--GASTVQLY----D 407
 473 AGAFTG--GH---GDALKYVTGTFDSEIPOCSLCEKHTRPVCAHTTVERLR----D 521
 408 ATRGWGLIDHNLVAGDVASIGSHLVRARVPSRPRENGWLVPGNSGEHWRGWLPHEAMP 467
 522 RFGQATRQPIGVFTKNSQY----SDCDPIGNYAP----LILRKPGDQTE 564
 468 R--VDPGGGIIVTANNRVYADHDPEYLCTDCHP----PYRAERIMRLVANPAFAVD 519
 565 AAKATHQDTYRATLERFLIDEQBERLLDRGA--PCSSBEGUSSVY----D 608
 520 DAAATHDLSP----HVGLLRRREALGARDSAAEGLRQMLVWGRMDAASEVASA 574
 609 HPTFRLDTL--RARIEOPTTQ----FMKVLTVERDYKIREG--- 645
 575 YNFAFRALTRLVDRGSLQASHPPAAVAFVSPQGQWAWPFLURDDAGMLKGMSW 634
 646 --LSEA----THSMALTEDPDYSCAFCP 666
 635 DQALSEALSVASQNLTGRSMGEHRFRTHPLATOFPAWAGLLNP 679

RESULT 15
 JS-08-633-760-48
 ; Sequence 48. Application US/086333760
 ; Patent No. 5804429
 ; GENERAL INFORMATION:
 ; APPLICANT: NIWA, MINEO
 ; APPLICANT: SATTO, YOSHIMASA
 ; APPLICANT: FUJIMURA, TAKAO
 ; APPLICANT: ISHII, YOSHINORI
 ; APPLICANT: NOUCHI, YUJI
 ; TITLE OF INVENTION: A NEW CEPHALOSPORIN C ACYLASE
 ; NUMBER OF SEQUENCES: 64
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
 ; ADDRESSEE: P.C.
 ; STREET: 1755 JEFFERSON DAVIS HIGHWAY, SUITE 400
 ; CITY: ARLINGTON
 ; STATE: VIRGINIA
 ; COUNTRY: USA
 ; ZIP: 22202
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC Compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/633,760
 ; FILING DATE: 01-MAY-1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: OBLON, NORMAN F.
 ; REGISTRATION NUMBER: 24, 618
 ; REFERENCE/DOCKET NUMBER: 18-929-0 PCT
 ; TELEPHONE: (703) 413-3000
 ; TELEFAX: (703) 413-2220
 ; TELEX: 248855 OPAT UR
 ; INFORMATION FOR SEQ ID NO: 48:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 774 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein

JS-08-633-760-48

Query Match 1.7%; Score 105; DB 1; Length 774;
 Best Local Similarity 20.7%; Pred. No. 0.5;
 Matches 146; Conservative 76; Mismatches 241; Indels 242; Gaps 37;

Qy 109 LTR--ACNAARERFGFSRCQQPPVDAVETTGAE-ICTR---LGLEPENTILYVVTAL 161
 Db 70 LTRKLAGRAEWLG--AEAAEADILVRRIGMEXVCRRAFEALGVEARD----M 117
 Qy 162 FRAEVFMCNVFIHYGGDILIVHNGDVIIRFLFPPYQLFMDPDRNVLYPDPNTNH----R 216
 Db 118 LRAYAVGYNFLASGA----PL-PYEVGLGAE--PEPEPMWSIAYNR 159
 Qy 217 SIGEGFVTPPYNTGLCHLHDCTVAPMAVALRTVNTAV----ARGAAHLAEDNEH 272
 Db 160 RLG----LMGSTMWKLWRMLALPUVGAANALKLRYDDGGRD 197
 Qy 273 AV-LPPDITYTYFQSSSSGTITARGARRNDVNSTSCKPSGGFER----RLAS---IM 322
 Db 198 LCLIPPAEARMLEADLATLPDVDLILRANGGDSDAAGSGSNNTAWAAGRATGRPL 257
 Qy 323 AADTAHLAEVIFTNTGIEYETTDIKWPMTGMECTLPRUNALGSYTARYAVGIVAMVFS 382
 Db 258 AGDPHRYFEI--PGIYAQHHLACDREDFM-IGL--TVP----GVPEPFHFA 298
 Qy 383 PNSA----LITTEVEDSGME----AKDGGPGESENRF 412
 Db 299 HNGKVAYCVAHFMDIHDLYQEFGAEGRFGRNDFEPVAWSRDIATRGADREFDIV 358
 Qy 413 YQAGPHIAANPQTDGHWUQSQTGSNTBEFSYDYLALICGFAQILLFLYLERCD 472
 Db 359 KTPHGPVIAGD--RDGAALTLSQAEFTDSLDCTEMP--GASTVQLY----D 407
 Qy 473 AGAFTG--GH---GDALKYVTGTFDSEIPOCSLCEKTRPVCAHTTVERLR----QRMP 521
 Db 408 ATRGWGLIDHNLVAGDVAGSIGHLYRARVERSPRENGWLYPGNSGEHEWMGWIPHEAMP 467
 Qy 522 RFGQATRQPIGVFTKNSQY----SDCDPIGNYAP----LILRKPGDOTE 564
 Db 468 R--VDPGGGIIVTANNRVYADDHDPEYLCTDCHP----PYRAERIMRLVANPAFAVD 519
 Qy 565 AXATMOTTYRATLERFLIDEQERLLDRGA--PCSSEGLSSVY----D 608
 Db 520 DARIAHADTLSP----HVGLLRRREALGRRDAAEGURQMLYAWDGMDAASEVASA 574
 Qy 609 HPTFRLDTL--RARIEOPTTQ----FMKVLTVERDYKIREG-- 645
 Db 575 YNFAFRALTRLVDRGSLQASHPPAAVAFVSPQGQWAWPFLURDDAGMLKGWSW 634
 Qy 646 --LSEA----D 666
 Db 635 DQALSEALSVASQNLTGRSMGEHRFRTHPLATOFPAMAGLLNP 679

Search completed: January 30, 2004, 13:18:32
 Job time : 27 sec

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

JM protein - protein search, using sw model

run on: January 30, 2004, 13:10:12 ; Search time 46 Seconds

(without alignments)

678.637 million cell updates/sec

Title: US-09-769-699-2

perfect score: 6294

sequence: 1 MENTQKTVTVPGLGXVVA..... DELFDLGSIBIKHGNITMEM 1203

scoring table: BLOSUM62

Gappx 10.0 , Gapext 0.5

searched: 830525 seqs, 258052604 residues

total number of hits satisfying chosen parameters: 8310525

minimum DB seq length: 0

maximum DB seq length: 2000000000

post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rabbit:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_viridis:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query	Match	Length	DB	ID	Description
1	4627	73.5	1194	12	Q9E1Y7		Q9E1Y7 cercopithec
2	3583	56.9	1208	12	Q9273		Q9273 equine herb
3	3371.5	53.6	1203	12	Q89549		Q89549 bovine herb
4	3327.5	52.9	1177	12	Q92611		Q92611 pseudorabie
5	3012	47.9	1197	12	Q69101		Q69101 herpes simp
6	2796.5	44.4	1190	12	Q9P252		Q9P252 turkey herb
7	2793.5	44.4	1190	12	Q9CTB9		Q9CTB9 marek's dis
8	2788	44.3	1191	12	Q9EP0		Q9EP0 turkey herb
9	2781.5	44.2	1190	12	Q9DPQ8		Q9DPQ8 meleagrid h
10	2764.5	43.9	1190	12	Q9E1G3		Q9E1G3 meleagrid h
11	2710	43.1	1191	12	Q9IBU6		Q9IBU6 turkey herb
12	1445	23.0	999	12	Q9QH63		Q9QH63 gallid herb
13	683.5	10.9	1132	12	Q40913		Q40913 kapozi's sa
14	675	10.7	1133	12	P88904		P88904 kapozi's sa
15	652.5	10.4	1103	12	Q99226		Q99226 murid herpe
16	651.5	10.4	1103	12	O41928		O41928 murid herpe

09/769,699 22-04
 Search Notes
 SEQUENCE FROM N.A.
 RA Gray W.L., Starnes H.B., White M.W., Ashburn C.V., Mahalingam R.;
 RT "Complete Sequence of the Simian Varicella Virus Genome."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP27348; AACG72021;
 DR InterPro; IPR000635; Viral DNA binding protein.
 DR PF00747; viral DNA bind.
 SQ SEQUENCE: EBA7F1C841965897 CRC64;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gray W.L., Starnes H.B., White M.W., Ashburn C.V., Mahalingam R.;
 RT "Complete Sequence of the Simian Varicella Virus Genome."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP27348; AACG72021;
 DR InterPro; IPR000635; Viral DNA bind.
 SQ SEQUENCE: _131768 MW; EBA7F1C841965897 CRC64;

RESULT 1
 ID Q9E1Y7 PRELIMINARY;
 AC Q9E1Y7; PRT; 1194 AA.
 DT 01-MAR-2001 (T-EMBL; 16, Created)
 DT 01-MAR-2001 (T-EMBL; 16, Last sequence update)
 DE 01-DEC-2001 (T-EMBL; 19, Last annotation update)
 OS Cercopithecine herpesvirus 7.
 OC Viruses; ssDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirus; Varicellovirus.
 OX NCBI Taxid:35245;

[1]

SEQUENCE FROM N.A.
 RA Gray W.L., Starnes H.B., White M.W., Ashburn C.V., Mahalingam R.;
 RT "Complete Sequence of the Simian Varicella Virus Genome."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP27348; AACG72021;
 DR InterPro; IPR000635; Viral DNA bind.
 SQ SEQUENCE: _131768 MW; EBA7F1C841965897 CRC64;

Query Match Similarity 73.5%; Score 4627; DB 12; Length 1194;
 Best Local Similarity 71.8%; Pred. No. 0;
 Matches 867; Conservative 143; Mismatches 178; Indels 20; Gaps 8;
 SEQUENCE: MENTQKTVTVPGLGXVVA..... DELFDLGSIBIKHGNITMEM 1203
 QY 1 MENTQKTVTVPGLGXVVA..... DELFDLGSIBIKHGNITMEM 1203
 DB 1 MEATQKTVTVPGLGXVVA..... DELFDLGSIBIKHGNITMEM 1203
 DR 1 GFSRCQOPPVGAVETGAETRGLPTEPNTLYLVVTALEAVFMNCVNLHYGGDI 180
 DR InterPro; IPR000635; Viral DNA bind.
 SQ SEQUENCE: _131768 MW; EBA7F1C841965897 CRC64;

181 VHINHGDPVIRPLPFVQLPFMDVNRLVPDFFNTHERSIGEGFYPTPFYNTGLHLHDC 240

Page 2

				PRELIMINARY;
039273	VRGRATVTRIPFLPIHFMFDTNRLLDPPNATORSISGENFTYPTPLFGTNCLLYDC	240	ID: O39273	ERT; 1208 AA.
241	VIAPMAYALRVNNTAVARGAAHLAEDENHEGAVLPPDITYFOSSSSGTTARGARRN	300	AC: O39273	
241	VISPIASLRNTRVAVARGAAHLAEDENHEGAVLPSDTTFSFOSSNGSYNPRLEPRON	300	DT: 01-JAN-1998 (TMBMLR); 05-Created	
241	VISPIASLRNTRVAVARGAAHLAEDENHEGAVLPSDTTFSFOSSNGSYNPRLEPRON	300	DT: 01-JAN-1998 (TMBMLR); 05-Last sequence update	
301	DVNSTSKSPSGFERRASIMADTAHLAEVTFNTGTYEPTDKEWPMIGMGTLP	360	DT: 01-DEC-2001 (TMBMLR); 19-Last annotation update	
301	---TSKSQT-GFERRASIMADTAQAEVFSGCVDEA.PSDTEWPMULGTONVTE	355	DS: Counterpart of HSV-1 gene UL29 and VZV gene 29.	
301	361 RINALGSYSTARVGVIGAMVFSNSALYLVTEVDSDMTEAKDGP3G5FNSRNYQFAEGPHL	420	GN: 31.	
356	RNLNLSYVARIETGVAAVMSNSNVYLTEIDGVSEGDGGP3FSNRYQFAEGPHL	415	OS: Equine herpesvirus 4.	
421	AANPQTDDEGHVLLSSQSTGSSNTESFSDYLAJCIGFAPLLFLYLERCDAGATGGH	480	OC: Viruses; dDNA viruses, no RNA stage; Herpesviridae	
416	AANPQVDEDGCVVFSGNGSVNTFNVDLALVGFSOLLARLFLYLERCDAGTGGH	475	OC: Alphaherpesvirinae; Varicellovirus.	
481	GDALKYTGTTGTFDSEIPCSLCBKHTRPYCAHTHTYHRLQORMPFQAOAQPICVFEGTNSQ	540	OX: NCBI TaxID:10337;	
476	SDALQYMTSTMNDIPCSLCERTRHLCVHTTIOQLKORMPRGPSPRQPIGEFGTNSQ	535	RN: [REDACTED]	
541	YSDCDCPLGNYAPVLLRKPGDQTEAATAKTMQDTYRATERLFLDEQEFLRDRGAPSSSE	600	RP: SEQUENCE FROM N. A.	
536	YSDCDCPLGNYAPVLLRKPGDQTEAATAKTMQDTYRATERLFLDEQEFLRDRGAPSSSE	595	RC: STRAIN=NS80567;	
601	GSSVIVDHPFTPLDTLRALEBOTQTQFMVLTVEVDYKIREGLSATHMALLDPY	660	PA: MEDFORD E. A.; Watson M. S.; Perry J.; Cullinane A. A.	
596	GLNSTIDHATFPLRALLEVKTRVQIIEOPKVLTVEVDYKIREGLTPEATHSLSITADPY	655	RT: The DNA sequence of equine herpesvirus 4.";	
661	SGAFCPITNLVYTHLAVVDQIALSOCHCVFYGOQVEGRNPNQFOEVLRERFDLFGN	720	RL: J. Gen. Virol. 79:1137-11203 (1998).	
656	SGSICPITYFLVTRNLIQDILSQHQGVTHQGOQVEGRNPNQFOEVLRERFDLFGN	715	RN: [REDACTED]	
721	GFSRSTSITVTLSEGPSAQNPLGQDAPAGRFTGDLARYSVEVIRLDRVNRVYFSGN	780	DR: [REDACTED]	
716	GFSRSTSITVTLSEGPSAQNPLGQDAPAGRFTGDLARYSVEVIRLDRVNRVYFSGN	774	QY: 1 MANTQXTVTVPGLGMYAACVEDDLEESFLALARSTDSD	
781	CINLSEEARLVLGASAYQREKRVDMHLGALGELLKQFHGGLPFGMPNPKSNPQW	840	DD: 1 MESAPKTVSLPSPLSPLGTVYA.CNTNFNETEAATLMARSIDS	
775	CINLSEEARLVLGASAYQREKRVDMHLGALGELLKQFHGGLPFGMPNPKSNPQW	834	QY: 1 MIAVSVSARTTOLLAGACITLKTTSHYPSVTPHGRKRVLF	
841	FWTLLQRNQMPADKLTHEBTIAAVKRFTEEVAINFNLPTCIGELAQYMANLILK	900	DD: 1 NTAVVAGSKTKTOLGGAGGITLKITPSPHTPNAFVYEGGSVFGAA	
835	FWALLQRNQMPAINTSSEINTIAAKRFTEEVAINFNLPTCIGELAQYMANLILK	894	QY: 121 GFSRCCQGPVDQDAVETGAEICTRLGIEPENTILYVTAFL	
901	YCDHSQTLINTLTSITGTCARRPRDPSSTLHWIKRDVTSAADIEQAKALLETEL	959	DB: 121 GSPPFSSPPVDAVETGAEICTRLGIEPENTILYVTAFL	
895	YCDHSQYINTLTAUTIOSKRPNPSSALHWDKEITPTDLETHARLIRNTQSPYM	954	QY: 181 VHNGDGYVIRTEFLPQLEMPVYNRLVDPDPTHRSIGEGHH	
960	WTAFTAFTSHTLYRAMNOCPMVGIGSISKYEGAGNNRVOAGNWGLNGGKNCPLFTF	1019	DD: 181 VTHQGPAVKTIEIYPOLYMPVNRNLAEPFSKRSIGDEEF	
955	WISSFMSINTLVRAMNQRPVNLGSIKYGAGNNRVOAGNWGLNGGKNCPLFTF	1014	QY: 241 VAPMAVALVIRNTAVARGAHHLADENHECAVLPFDITTH	
1020	DRPFRPITACPRGGTICPVYTGSSGMRTEITLSPVQRTGIVSGAMMQLAIYATVURANGA	1079	DB: 241 VCPAAVALVIRNLDTGARGAHHLADENHECAVLPFDITTH	
1015	DRTRFVTAICPRGGTICPVYTGSSGMRTEITLSPVQRTGIVSGAMMQLAIYATVURANGA	1074	QY: 299 RDVNSTSKPSISGGFERRASIMAADTAHLNEVENTGIVE	
1080	RAQMADFDDWLSLTDDEFIARDLLELHQIQLETPVVEGALEYKIL-DEKTAGDG	1138	DB: 359 RBRLEAGAYAQRSLSGINGANFSANSVLTMEVDDGPGADQ	
1075	RAEMGFDDWLSLTDDEFIARDLLELHQIQLETPVVEGALEYKIL-DEKTAGDG	1134	QY: 419 HIAANPOTDROBVLH-S---SQQGCSNNTFSDYLAIC	
1139	ETPNLAFNFDSCPSPHDTSNTVNISSGNISSSTVP--GLKRPPBDEDFLGSPIK	1195	Db: 418 YAGNPQDIDKDRVLOHTADQAAPIGNQSPSLYLAAC	
1135	SGCGNIAATNFDTCNDSS-----NTEENHAPPSSYOPLGSCGKRPDDDLTDMSSLPK	1186	QY: 359 LPRNALGSYTRAVGTTGAMFESPNSALYTTEVEDSGMTR	
1196	HGNMEM 1203		DB: 478 AGFGFGENNETDRLYRANTLLEBVPCCCLCTPAPRCAHTT	
1187	M-SLMDM 1193		QY: 532 GFGTGTSQYSQSCDPLCNYAPLILKXPGDQTEAAATQMDT	

592 DRGA-----PCSSSEGLSSVIVDHPPTERRILDTRARIEQTTOQFMKVLYVERDYKIREGL 646
 593 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 597 DKEELAHAGTCSA---TGIVVKDQASFPNLLSTIKDITEGAAQFVRTLVEVRDFKIREGL 654

 647 SEATHSMALTFDPYSGCATCIPINFLVERTHLAVQDIALSQCCHCVTYGQVEBGRFRNQF 706
 648 ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
 655 ADANTHMSLSDPSSSFCEPVTSFLSRTIFAVLQDLVLSSQCHCLTYGQSYTEGRFRNQF 714

 707 QPVURRFVDLFLNGFISTRSTVILSSEGVPASAPNPTLGQDAPAGRTFGDGLARYSVQE 766
 708 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 715 QPVURRFELDMLANGFPIAKTVTIVTUSGSVTAAPNLTPSSRBPPTCDYGDMMARSMEVL 774

 767 RDIRVKRNRYVFSGNCTNLSEAAARPLVGLASAYORCERKVDMLHGALGFLLKQFGLLFFP 826
 768 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 775 RDLEIKRNRLPSNGANNSSEAAARVAGMASAARPRPERGSNTLNAGVFLYKQFKRVLFPP 834

 827 RGMPNSKSPNPQWFWTLQRNQMPADKLTHEETTTIAAVKRFETEYAAAMFINNPPTCI 886
 828 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 835 RGHPGIDTPNPKQFWTLQRNQMPARLISKEDIETTAIKRFSHBSYSAINFINTPNNI 894

 887 GELAQFYMANLILKXCDISQYLINTLTSITGARRPDPSVLUHWIRKDVTSAADLETQA 946
 888 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 895 GELAQFYFANLVKXCDISQYFINGLIAIVGSRPRPDPAAVLAWINTRINGSADVEPAA 954

 947 KALLEKTENLPELTWTAFTSTHLVRAAMNQRPVWVIGTISIXXHGAGANNRVRFOAGNWSG 1006
 948 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 955 QEVVQLGNNPAAWTGTFASTINNVRWYMDQRPAVIGLSISYNGSAGNNRVRFOAGNWG 1014

 1007 LNGGKQVNCSEPLFPTDRTRRFLIAQPRGGFICPVTGPGSSGNRENTLSSQVRGJIVSGGAMYQ 10666
 1008 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 1015 LNGGKQVNCPLMAFTRREVLAQPRGVGETCEAGFGNGARENTLSSQQRISIVSDGPVWQ 1074

 1067 LARYATVURAVGARAQHNAFDWMSLSDDEFLARDDELLHQIIQTLTPTVGEALEAV 1166
 1068 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 1075 TAVSVVLTALGARTOHALDDWGLVYDDEFLAASDALNAAVDQF-GEWSVEAQDMI 11133

 1127 KILLDEBT -----TAGDGEGTPNTIAFNEDSCEPSSHDTISNVNLTISGSNISGTVPGKRPP 1181
 1128 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 1134 RTMPLAQTMNGVUTSGDG -----AFDGFACVGDNQSSTFNMGPA--SSAAPAGQKRHF 11855

 1182 EDDPLFLDSLGIPKXGHNTMEM 1203
 1186 PDDFLDGAPPERKXSGLTFDM 1207

PI	Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
RN	[4]
RC	SEQUENCE FROM N.A.
RA	Schwyzer M., Submitted (BSR-1997) to the EMBL/GenBank/DDBJ databases.
RL	EMBL: 278205; CAB01596.1;
DR	EMBL: X54677; CAB01336.1;
DR	EMBL: AU004801; CAB0104.1;
DR	InterPro: IPR000355; Viral DNA bind.
DR	Pfam: PF00747; viral-DNA-bp_1.
SQ	SEQUENCE : 1203 AA; 127409 MW; 8299D64966A9654F CRC64;
Qy	Query Match 53.6%; Score 3371.5; DB 12; Length 8.1e-269;
DR	Best Local Similarity 53.6%; Pred. No. 197; Mismatches 340; Indels 6512; Conservative 197; Matches 197;
Qy	: 1 MENTQRTTVPGLGYYACRVDELDIIEISFAARSDSIALDPLPPVYV
DR	: 1 MDAAKTVALAPBAGFTVCDARLDLKLALIAARSBDSSLAVLP
Qy	: 61 SLAVYSGARTTGLAGAGITLKLTSHFSVPSVYEHGKRVLPSSAAPI
DR	: 61 NVAVAGTTTGTGAGGTTAKLPHPSYHPPNPFERGGBLRSTAAPP
Qy	: 121 GFSRCQGGPVGDGAETTGAECIPRLGLSPENTILYVVTALFEAEAV
DR	: 121 GFSAGAGGPVGDGAETTGAECIPRLGLSPENTILYVVTALFEAEAV
Qy	: 181 VHNHGDTTRIPLEPPVQLEMPDYNRLVDPFTHRSIGEGFYPTP
DR	: 181 VOVGAGEAVRVPVLYPVQLYMPDVTRVNEKEPNARQRAQEOLAYPRP
Qy	: 241 VIAPMAVAlVRVNTAVARGAAHLAFDNENHEGAVLPDPTTYTFOSS
DR	: 241 VLGPAANAVLRVRLDAVARGAAHLAFDDSHEGAVLPDPTTYTFOSS
Qy	: 301 DVNSTSKSPSGGFERRASIMADDTAHLAEEVIFNTGGYYEETPDIDKH
DR	: 300 ADPAAKAAPGGYERLAVSNMADTAASVIEAMSTSFEDDEDACVDD
Qy	: 361 RIVALGSGTARYAVIGANVFSNSALLTVEEDSGTMKDDQGPGP
DR	: 360 KLDALGAVGRLLAGLVGMFSSNSVLIIMTEVDGGAAADAKOGA-AAA
Qy	: 421 AANQTDGHDYLISSQSTG-----SSITEFSYDYLAIICGEGAPLIA
DR	: 419 AGNPRCDDKGPKLPQTGGPAYSINGAQEFALDHLLAACGFDPQLLW
Qy	: 475 AFGNGHG-DALKYVITGTPDSEIPLCSLCKHTRPYCAGTIVHLRORM
DR	: 479 AFAGRNDALKYVASTLGGTNTGASYSAIRRPG-ADENARSIMGATVAAENV
Qy	: 514 GAPSSGEISSLVVDHPFRRLLDTLRARIOTTOQKVLYBTDRDYM
DR	: 538 DASAAQ-LERATDHAISFRGALLAIQTVEQATEAFTGQVTDRDPK
Qy	: 594 ALTFDPYGAFCITNFVLRKTHLAVYQDLAISQCHCFYQGOVEGRN
DR	: 599 FGTMNSAASDCDVGNTAQSYSAIRRPG
Qy	: 654 GAPSSGEISSLVVDHPFRRLLDTLRARIOTTOQKVLYBTDRDYM
DR	: 657 SLDLDPYATCPATAFRRSLTAWQDLAUSQCHGFCGQPVDGMRN
Qy	: 714 FVLDENGPRISPSISITLSEGPVSAPNPTLQGDAPARTEDDQDAR
DR	: 717 FMQLINGCGFLTRITVTVLAEARAVERPAAQAQTEPPARDMDDISK
Qy	: 774 RVFSGNCNTNLSEAARALVLGLASAYQQEK-RVDMLGALGELJKQO
DR	: 777 RMESAGNMMSZEARAUGLACAYQDESGGNILSOSPLGLVVKQ

833 SKSEPNPQFWTLIQRNOMPDKLTHEETTIAAKRKRTEEEYAAININFNLPTCIGELAQF 892
 Qy :|||:|||:|||:|||:|||:|||:|||:|||:
 Db 837 SPPNPQFWTLIQRNOMPARIISKEDLETIAKRFSDATAAINVNLTNLGNVAELAQF 896
 Qy :|||:|||:|||:|||:|||:|||:|||:|||:
 Db 893 YMANLILKYCDH5QYLINTLTSITGARRPRDPSVLHWIRKDVTSAIDIETQAKALLEK 952
 Qy :|||:|||:|||:|||:|||:|||:|||:|||:
 Db 897 YMANLVLRHCDH5QFYINGLTAVAGARRPRDAAWHWTERRIADAPDAAAEBVLRE 956
 Qy :|||:|||:|||:|||:|||:|||:|||:
 953 TEULPELMTTAFTSTHLYRAAMNORMPVYLGI-SISKHGAGACNNRYFOAENNSGNGGHN 1012
 Db 957 AEARPEWVAGTFFASNLRSYNTMTRPAVLGLISIKNGSAGNRYFOAENNSGNGGHN 1016
 Qy :|||:|||:|||:|||:|||:|||:
 1013 VCPCLFTFDTRRDLIACPRGGFCPTGPSSGRRTTLSQVRGII-VSGCAMVOLAIYA 1011
 Db 1017 VCPCLCFKTRRVYLACPRAGEFCATAGAGGANDNTLVEALEIINGSGALAAQTAVYG 1016
 Qy :|||:|||:|||:|||:|||:
 1012 TYTRAVGARAQHAFDDWLSLTDEFFLARDLLEBLHDQIOTLETP-WTVEGALFAVKIL 1139
 Db 1077 AVLQLGRVEMHDDWAAVDEPFAQSMTLVE-RAAARPQGWSPGAAMLREL 1133
 Qy :|||:|||:|||:
 1130 DEKITAGDGETPTPLNAFNDFSCPSHDTTSNYLN--ISGSNTGSSTVPGKLKRPPDDELE 1137
 Db 1134 -ELEAAAE- PFGGAADFCAASAAPPDATIAFGGPVAG--AGPAPGAKR- PUDLELF 1187
 Qy :|||:
 1188 DLSGIPIKHGNITMEM 1203
 Db 1188 EM-GVPDRKRPAUTDM 1202

RESULT 4

092611 ID 092611 PRELIMINARY; PRT; 1177 AA.

AC 092611;

DT 01-NOV-1998 (TREMBLrel. 08, Created)

DT 01-OCT-1998 (TREMBLrel. 08, Last sequence update)

DE DNA-binding protein.

GN DEB_P.

OS Pseudorabies virus.

OC dsDNA viruses, no RNA stage; Herpesviridae;

OC Alphaherpesvirinae; Varicellovirus.

RN [1]

RP STRAIN-TNL;

RX MEDLINE=98455382; PubMed=9784061;

RA Wu S.-L.; Hsiang C.-Y.; Ho T.-Y.; Chang T.-J.;

RT "Identification, expression, and characterization of the pseudorabies

RT virus DNA-binding protein gene and gene product."

RL Res. 56:1-9 (1998).

EMBL: U80905; AAC63429.1;

DR InterPro; IPR000635; Viral_DNA_bind.

DR Pfam; PF00747; viral_DNA_bp; 1-

KW SEQUENCE FROM N.A.

DB STRAIN-TNL;

DB SEQUENCE 1177 AA; 125408 MW; BA87AF9CC961707 CRC64;

SQ SEQUENCE 1177 AA; Score 3327.5; DB 12; Length 1177;

Best Local Similarity 53.5%; Pred. No. 3.4e-265;

Matches 655; Conservation 182; Mismatches 318; Indels 69; Gaps 17;

Qy 1 MENTOKTYVPTPGLGYYYACVVEDLDEETSFLAARSTDSDLALLPMENLTKEKTFPS 60
 Db 1 MEAAKAKTVVRAPLGTGVVTPBEAERDLALVALVRSADDAAAYAPLVRGLTVEAGFAG 60Qy 61 SLAVUSGARTGAGAGITLKUTSHPKSYTVEFGKAVTDSAPNLTTRACMAREF 120
 Db 61 HVAVVAAGARTGI-GGGHTLKTUPNTHPNTFFERGDCYPSSAPALSRAACEARAF 119Qy 121 GFBCRCQGPVVDGAVETTGAEICTRGLGEPENTILYLVVTALEKAVFMCNVFLHYGLD 180
 Db 120 GFSEAYR-TFVVDNEBETTGAEVARLGAAADAHAYLVADKPEAYLCHNMFYNGGAGT 179

Qy 181 VHINHGDVIRIPLFPVOLFPMDFNRVLPDFPNTHERSITGGFVYPTPFYNTGLCHIDC 240

Db 179 VSINGHEAWRVPFLYVPLFMPDVNLVADPPFKARNESISEEFVYPPFPFFNPLCMLLHG 238
 Qy :|||:|||:|||:|||:|||:|||:
 Db 241 VIAPMAVLRVNTAVARGAHLAFTENHESAVLFPDITYFEGSSSSCTTARGARRN 300
 Qy :|||:|||:|||:|||:
 Db 239 VLGDAAVTRVNLNTRVPLFEEORRGDDAA----- 292
 Qy :|||:|||:|||:
 301 DVNSTSKSPSGSGFERLIASTIMADPTALHAVTPTNGTIVETPTDIKEWMP----- 352
 Db 293 -----AGLERERMASMSADALSLTEALVAGVYDEPPLAIDDWPTLSEAAGAKD 342
 Qy :|||:|||:|||:
 353 ---IGMEGLTPRLNLAGSYTARVAGVIGANNEFSNPSALVITEVEDSGMTAEKOQGPGPS 408
 Db 343 GALAASAGAPVPSQOAAALGAYVSRAGIGVGNLVSNSVVLVITEVDAGAADARKEGAGPS 402
 Qy :|||:|||:
 409 FAFFYQFAGPHLANPQODRDRHVL---SQQSTSNTESVTDIALCFFGABPLARL 464
 Db 403 ENFYYQTAAPYLAQNPQTDKGIVRULLHTASSQATAPGPNHDEPMHVMACFCPCOLLARY 462
 Qy :|||:|||:
 465 LFYLCRDAGAFTG-GEIGDALCXVTCNFDSPIPSCLCEKTRPVCAHTYTHRLRORMPRE 523
 Db 463 LPFLERCDAGTGFGRGIVDQYDARVSPACHTHILRLR-SRA 521
 Qy :|||:|||:
 524 GOTROPICGVFTMNNSYSDPGLNAYPAVYUILLREGDCTEAAKATMDYTRATEBLFLI 583
 Db 522 SARRSPNGVFTMNNSYSDCIVLGNAYASTALKRND-SEPPKALMQDTRAAYDRLLA 580
 Qy :|||:|||:
 584 DLQEERLLDRAPIGCSSEGLSSVINDPTEERILDLVTPARIEQTTQFMKVWVETPDYK 643
 Db 581 DIVAGR -----IGEVTVDEAGFREALRAIRDYTVQADFRVTLVETRDFKL 628
 Qy :|||:|||:
 644 EGLESIAETHSMALTEDPSGATCOPINFLVKEATHLAVQDIALSQCHCUPVQGQTEGRNR 703
 Db 629 DALYDANTHMSLSDPSGALCPANSFELARETLLAVLQDIALSQCQGVHVGQTEGRNR 688
 Qy :|||:|||:
 704 NOQPVLRPFRYDLFNGEFTSRSITVTLISGPVSAPNPLIQDAPAGRDFGDLYRSV 763
 Db 689 NOQPVLRPFRYDMLNGGFVTAKTVNTLQAG-IVADPDTKGSAPPADHDGDLYRSV 747
 Qy :|||:|||:
 764 EVRIDYRKVVPSGNCNTNSEAARLYVLASAYQDQERKRVDMLEGAFLIKQFHGL 823
 Db 748 EVLRELVKSRMFSTSGSISDARARVAGLAGAYQRPDDAVDINGPFLKQKHAT 807
 Qy :|||:|||:
 824 LFRGMPPNSPSPNPONFWTLQRMNMPADLTHETITLAIVRKFTEETAAINFNLNP 883
 Db 808 LFRGKPKQQSISNPQFWTLQRNQLPARLTKDQETAAKRFSDGAINVNLTQ 867
 Qy :|||:|||:
 884 TCIGELQFYMNLILKYCDQSVLINTLISITGARRPDPSSVLEWTRKDVTSAADIE 943
 Db 868 GTAAELAQFYLNLIYRCDKQFFINSLGITQSKRDPDAAMNTYRPLADAADAE 927
 Qy :|||:|||:
 944 TQKALLEKTNLPENNTATSTHVRAAMNQPNVNLGISISYKGAAGANRNFQAGN 1003
 Db 928 RAAREVLDAPR- -DWVATYSSHULRSVMAPIWVGLGVSKYHMGAGNNFVQAGN 985
 Qy :|||:|||:
 1004 WSSELNGGGKHHVCPMLVDFTRAPIVACPRVGETCSOTGGAGLHDKSIGHVKTLADGSP 1003
 Db 986 WSSELNGGGKHHVCPMLVDFTRAPIVACPRVGETCSOTGGAGLHDKSIGHVKTLADGSP 1005
 Qy :|||:|||:
 1064 MYQIAATIVVAGRAQAGNAFDDMLSLTDEFLRLDQIOTLTPTVVEGAL 1123
 Db 1046 LYQTAIAAVLGLARTQHEDFLALAEINGRVADE-DGRMSVEAAA 1104
 Qy :|||:|||:
 1124 EAKKILDEKTTAGDGETGEPPTNIAFNFDSCSEPSHDTTSNVLNISGSWNIS-GSTVPP--GIKR 1179
 Db 1105 ELVRLDQIOTGAGGGBET-AFDGAGCAGACD-----AGAGLAPASLAPALGGRR 1154
 Qy :|||:|||:
 1180 PPFDELFDLSLPIKHNNTTEM 1203
 Db 1155 PPFDELFDLSLPIKHNNTTEM 1176

Q69101	PRELIMINARY;	PRT;	1197 AA.
C069101;			
IT 01-NOV-1996 (TREMBrel) 01; Created)			
IT 01-NOV-1996 (TREMBrel) 01; Last sequence update)			
IT 01-DEC-2001 (TREMBrel) 19; Last annotation update)			
E DNA binding Protein ICB8			
S Herpes simplex virus (type 2)			
S Viral DNA viruses, no RNA stage; Herpesviridae;			
C Alphaherpesvirinae; Simplexvirus;			
NCB_ TaxID=11310;			
P SEQUENCE FROM N.A.			
C STRAIN=KN;			
A Tchih Y, Tanaka S, Liu Y, Mori R;			
I Nucleotide sequence of the major DNA binding protein of herpes simplex virus type 2 and comparison with the type 1 counterpart. ;			
L Arch. Virol. 129:183-196 (1993).			
EMBL; D10558 BA01501.1;			
R InterPro: IP000635; Viral DNA bind.			
R Pfam: PF00741; viral DNA bind.1.			
Q SEQUENCE 1197 AA; 128470 MW; AA3A758865BPF CRC64;			
Query Match	47.9%	Score 3012; DB 12; Length 1197;	
Best Local Similarity	49.0%	Prd. No. 4e-239;	
Matches 595; Conservative 212; Mismatches 368; Indels 40; Gaps 16;			
Y 4 TQKTVTVPGLGKVY-ACRVEDDLEEISFLAARSTDSDLALLPLMRNITYKTFPTSS 61			
b 7 TTTVKVPGPMGVYGRACPRLDE-LSLSLRSRGDAADVAPLVGTVESSFEAN 64			
Y 62 LAVVSGARTTGAGAGITKLITSHPSVNFHGGHVLPSAANLTACNAERFG 121			
b 65 VAAVGSGTTGGLGCTAVSIKLMSPHSVSVYFHGGHLAPTAQNLTRCERARRHFG 124			
y 122 FSRQGPVDPGAVETTGAEICPRGLEBENTILYVTALEKAVEMCNVFLHYGGDIV 181			
b 125 FSDYAPRCDLKHETTDALCELRGLDPDRALLYVTEGFEAVCISNTFHGMDDKV 184			
Y 182 HINGDVTRIPLEFVQLMPDVNRLLVPPEFTNTHRSIGEGFVPTPYNTGICHJHDCCY 241			
b 185 TIGDAEVTRIPVPLQMMNPDSRVIAFPNCHRSGENFYPPLPNFLARLIFEA 244			
Y 242 IAPMVAVALVRVNNTAVARGAAAHAFDENHEGAVLPDTITYFQSSSSGTGTARGARRND 301			
o 245 VGPAAVALRARNTDAVAAAHHAFDENHEGALPAITTAFAESCG--KPGQGAR-- 299			
Y 302 VNISTSKPSPGGFERRLSIMADTAHAEVFNFTGYEPTDKWKPMTIGMESTLPR 361			
b 300 -DAGNKGPAAGQRLASMGDAALALESIVSMAYFDEPPDITTPWPLEQSTPAIR 357			
Y 362 LNAAGSYTARVAGTGANVFSPSALTYTEVEDSGMTBAKDGPGISFSNRYQFAPPHLA 421			
b 358 AGAVGAYLARAAGIVGAMESTNSALHTEVDDAGAPDZDHSK-PSFYRFLPVGTHVA 416			
Y 422 ANPOTDRDGHV-----SSQSTSNTESVSYDYLALICGFAPILLARLFYLERCDA 474			
b 417 ANPOLDREHVVPGYEGFRTAPIVGGQ-EFVGELAHLICGFSPALIARLFYLERCDDG 475			
Y 475 AFTGGHG-DALKVYTGDFSEIPOSLECKHTRPVCAHTVHLRQMRPREGQATROPIGV 533			
b 476 VIVRQEMPVFRYADSSQTDVONLCTFETTHACATTMLVRLRAHPKAASAQAIKV 535			
Y 534 FGTMNSQYSDCDPLGNKAPYLIRKPGDQTEAKATHQDTRATEFLFIDQEQLLDR 593			
b 536 FGTMNSQYSDCDVLYAFAFSALKR-ADGSENTRNTMQETYRAATERVMABLEAQYVQDQ 594			
Y 594 GAPSESEGGUSVVDHPFERRLTLRARIQQTQPMKVLYVBDYKIREGLSARTHSM 653			
b 595 AVPFLALGKLETINREALHTVNIKOLVDEVEQMRNLLEGKFKRGLDANHAM 654			
Y 654 ALTDPYSGAFCPTTNFLYKRTHLAVYQDALSQCHCVFGQVEGRNFRNQFQFVLR 713			
Db 655 SLSLDPYTCGCPCLQLARRSNLAVYQDALSQCHGVPGGSVEGRNFRNQFQFVLR 714			
Qy 714 FVDLNGGGFEGTTSRSTTUTSEG-EVSAPNPTLCODAPAGRITGDLARSVETVDRVK 772			
Db 715 WMDLNGFLGFSAKTLVTLVSEGAICASLTQATAPESSEPGDVARTLFFKEVRVK 774			
Qy 773 NRVYFSGNCTNLSAARALVGLASAYOROEKRDMLHGALGFLLKQFHGLFPRGMPPN 832			
Db 775 SRVLAGASANASAAKAVSLOSSAYQDPKDVLGPGFGLKQHAYFFNGRPGP 834			
Qy 833 SKSPNIPOWFWTLORNQMPADKLTHEETTIAAVKRFTEEYAAINFNLPPTCIGELAQF 892			
Db 835 SNOPNIPOWFWTLORNQNPALARLSDREDIATIKRPSLDSYGAINFNLAPINVSEAMY 894			
Qy 893 YMANJILXICDHSQLINLTSITGARRPRDFSSVHWIRDUTSAADIEQKALLEK 952			
Db 895 YMANQILRVCDDHSTYFINTLTAVTAGSRPSPQYAAAAPPQ--GGAGLEGARALMDS 951			
Qy 953 TENLPBLWTTAFTSTHLYRAAQNQRPVYVLSKCHGAAANRNVEQAGNNSGLNGK 1012			
Db 952 LDAPGANTSMPSACSONLRLPVMARPMVTLGDISKYGMAENDRVQAGNASSLQGRN 1011			
Qy 1013 VCPLEFTFDTRFTIACPRGGFICPVGTSSENRETTLSDOYRGIIYSGGMVQLAAT 1072			
Db 1012 ACPLLIFDTRKFVACPAGFPCAASSLGGAAEHSCQRGQITABGGAAVASSYFVA 1071			
Qy 1073 VRAVGAHQMYADDMLSLTDDFLADLEBLHDQIYQTLTPWTFEGAL--BAVKI 1128			
Db 1072 TVKSLGPRTQQLQEDWALLEDYLSEMMPFTTRALEGRGEWSSTDAALEVAHEAEAL 1131			
Qy 1129 LDEKTTAGDEGETPTLNAFNEDCEPSHDTNTLNIGSNSIGTSTVPLKRPEDDELFD 1188			
Db 1132 VCOLGAAGEA----VENFGDFGDEDQHAAASFGLAA-AGAAAGYARKPAFHGDPPFG 1183			
Qy 1189 LSGPPIKHNTMEM 1203			
Db 1184 -EGFPEKK-DLTDMD 1196			
RESULT 6			
Q9P252 PRELIMINARY;		PRT;	1190 AA.
ID Q9P252			
AC OQPZ52			
DT 01-MAY-2000 (TREMBrel. 13, Created)			
DT 01-JUN-2001 (TREMBrel. 13, Last sequence update)			
DB "Identical sequence to the Homolog of HSV-1 Glycoprotein.			
OS Turkey herpesvirus.			
OC Viruses: dsDNA viruses, no RNA stage: Herpesviridae;			
CC Alphaherpesvirinae; Marek's disease-like viruses.			
NCB_TaxID=10390;			
RN SEQUENCE FROM N. A.			
RP			
RC TRAIN=HP24;			
RA Kato K., Jang H., Izumiya Y., Cai J., Tashima Y., Miyazawa T.,			
RA Kai C., Mikami T.,			
RT "Identification and Transcriptional Analysis of the Marek's Disease Virus 2 Genes Homologous to the Glycoprotein B (UL27), the ICPI8.5 (UL28), and the Major DNA-binding Protein (UL29) Genes of Herpes Simplex Virus Type 1."			
RT EMBL; AB024711; BAA83153.1;			
DR InterPro; IPR000635; Viral_DNA_bind.			
DR Pfan; PF00747; Viral_DNA_bp; 1			
DR SEQUENCE 1190 AA; _130339 MW; D9291E2EEA7F59D_CRC64;			
Query Match 44.4%; Score 2796.5; DB 12; Length 1190;			
Best Local Similarity 45.3%; Pred. No. 2_6e-221;			
Matches 546; Conservative 235; Mismatches 396; Indels 29; Gaps 13;			
Qy 6 KTVTVPGLGYYACRVEDLDEEISPLAARSTSDSPLALLPLMRNLTVERTFTSSLAVV 65			

Db	5	KSYKITGGPIGYYYAAPTSSMPAEDLSIYAAKSNDCEDAILPLVSGLTVEADMNNVAAV	64	Qy	1139	ETPNLAINFDSCEPBDTTSVNLNSGSITVGLKXPPDEBELLGDISGIPKRGH-	1197
Qy	66	SGARTTGAGAGITLKLTTSHFPSPSIVVERGGKHVLSSAAPNLTACNARERFESRC	125	Db	1137	DGPT---TNYDALDENTERRAD----GPSI---LEPTIKRCPSD--VFDELPVFKRAP	1183
Db	65	AGTKTTGCGSTTLKVPTVPHGECIKRCTAQNLTACDLARGEYS-T	123	Qy	1198	NITMEM 1203	
Qy	126	QGPVDGAVETTGAICBICTRLGPEENTLYLVTAFLKEAVFMCNVFLHYGGLDVHNE	185	Db	1184	GLSYDM 1189	
Db	124	YSPVATSPETEQICDNLGNPQEIMLYLVTELEKAVLNCNSYLHGTGVVING	183				
Qy	186	GDTTRIPLEPPVOLMPDVNLRYEDPFNTHHRSTGEDEVYPTFFNTGLCHLIHDGVIAFP	245				
Db	184	VDRRIPLPHIIVFPDENRVASDPFSIKPRAUGEGALMPAFYNDLSCLRLHGVLSTA	243				
Db	246	AVALRVRVNTAVARGAHLAFDNENHEGAVLPPDITYTPQSSSGTTARGARDVNST	305				
Qy	244	AVGRVRVANDAARCIAHLSFDNHECTLLPADTAAFTPEAKRSQYAKKGKGMEL	303				
Db	306	SKPSPGGFERIALASTMAADTLHAEVTFNCIYEETPTDKWMPMFIGMEGLPRLNAL	365				
Qy	304	S---GGYERRTAVALMSADATSIENVIATSVYEEIPDVVKWPYCSPIGTYTRVEAL	359				
Db	366	GSTARAVGIGAVFSPNSNALLTEVEDSGMTEAD-GGGPSPNTRFYQAGPHLAAP	424				
Qy	360	SAYMARTAVGIGMVFESENNSNVYIYVYIYVYIYVYIYVYIYVYIYVYIYVYIYVYIY	419				
Db	425	QTRDGHYLS---SQTGSTNTMEFSYDYLALICGFCAPLARIYLFLYERCDAGFTGCH	480				
Qy	420	LIRDGRPIPGEELSATSASSEYSIDYLILACGPQOLLARFLYLEBCDGQSACH	477				
Db	481	G-DALKYTVGTGPSEIPCSLCHEKHTPVCATHTVERLORMPRFGCATROPIGEJGTNS	539				
Qy	480	DLTUVKFYSSAMADVPCDLCRASRUYCAHTKRELYRIPKFOYQMGAMGLGSMTN	539				
Qy	540	QYDCDPIGNYAAYPLILRKPDQTEAKATHQDTATRBLFLIEOBFLDGAPEPCS	599				
Db	540	NYCDNAIGSAYSFSTUKR--SEGTSRSYMDQTYKLTVERVMKALEGKLTZEDPAIM	597				
Db	600	EGLISSVIVDHPPFRRLDITLRARIBQTTOQMVKLVTRDYKIRGLESATHSALTFDP	659				
Qy	598	TBDALLRDGKSFMRALSTMANIESEAGQMLNLTIREYNTIREGLGATHTLSATEP	657				
Db	660	YSGAACPITNFLYKTHLAVYDLAUSQCHCVFQGQVEGRNFRNQFQPYLRREFDLEN	719				
Qy	658	YSSGICPVLSFSRRTIAYQDMAISQCSNMIMHGQOVEARNFRTOQAVLRRVDLON	717				
Db	720	CGFISTRTSLTSEGVASAPNPTLGQDAPAGRTEFGDALARSEVIRDRYKRVYVFQ	779				
Db	718	AGPITSKNITVLEDQISVPPSKSQHDLPLAVMEGLYXVTFEFIREKVKVMFVG	777				
Qy	780	NCTN-LSEAAARLVLGLASAYORQEKRVDMLHAGFLKQFHGFLFPQGMPPNSKSPNP	838				
Db	778	GSVTSVDTGSLAGTNEATORPAGMYHNGPLCFLARYHTPLFPAVKGPTPENA	83				
Qy	839	QWFWTIQLQRNOMPADLCTHEETIAAVKRPTTEYAAINFNLPPTCIGBLAQYMANLI	838				
Db	838	IWFWTIQLQRNLPAGLSKDAEENTSKPRTNTSADMTNINISPTCFGEGLAQYLYATTI	837				
Qy	899	LKPCDQSQYLNTLITIGTRPDSVLIWIKDVTSAADETQAALLETTENDE	938				
Db	898	LYCSHKFFINTISALVATESKPRPDMVLPWIEHILQGSDVAPAQOLLKNICDKE	957				
Qy	959	LWTAAFTSTHILVRAAQNQRPNMVLGLISISKYHGAGNNRUFQAGNWGSLNGKVNCPFLT	1018				
Db	958	AWCAAASSTNLVGPIMASKPTVWLTAVLVEGAEVNLKIDETTADG	1138				
Qy	1019	FDRTRPFIACBGGFICPTGPFSSNRETTLSOYRGTLVSGGAMYVOLATIAYVVRVAG	1078				
Db	1018	FNRTHRFVMACPGVFSBQGTFSSLKETLVDAVUSECZGAPAHVNMIAKONG	1077				
Qy	1079	ARAQHMAEDDNLSTDEELARDLEHDLQICLTFPTWVEGALENVILDETTADG	1138				
Db	1078	DRVQMEDDDDMEITDAYTSLIDLNKQ-VEGEGGNSVDAAILAKEMVNDISKALPL	1136				

Qy	1139	ETPNLAINFDSCEPBDTTSVNLNSGSITVGLKXPPDEBELLGDISGIPKRGH-	1197	Db	1137	DGPT---TNYDALDENTERRAD----GPSI---LEPTIKRCPSD--VFDELPVFKRAP	1183
Db	1137	PRELIMINARY:		Qy	1198	NITMEM 1203	
Db	1184	Q9QB9		Db	1184	GLSYDM 1189	
Db	246	Q9QB9		Db	246	AVLRTAVTAVGAAHLAFDENEGAVLDPDTYTHQSSSGTITARGARANDVST	305
Db	244	Q9QB9		Db	244	AVLRTAVTAVGAAHLAFDENEGAVLDPDTYTHQSSSGTITARGARANDVST	305
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	
Qy	1198	Q9QB9		Qy	1190	AA.	
Db	1184	Q9QB9		Db	1184	AA.	
Db	246	Q9QB9		Db	246	AA.	

306 SKPPSPGGFERRLASIMAAADTALHAEVIFNTGLYEEPTPDIKEWPKMFIGMEGTLPRLNAL 365
 Y |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 b 304 S---GGYERRTATSLMASDATISIENIATSYEESTPDIKPKWLYCSPGTYDVEAL 359
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Marek's disease-like viruses.
 NCB_ TaxID=10390;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M65;
 RX MEDLINE=10392152; PubMed=1033706;
 RA Tuinan E. R., Afonso C. L., Lu Z., Zsak L., Rock D.L., Kutish G.F.;
 RT "The genome of a very virulent Marek's disease virus.";
 J. Virol. 74:7980-7988 (2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M65;
 RA Tuinan E. R., Afonso C. L., Lu Z., Zsak L., Rock D.L., Kutish G.F.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
 DR EMBL: AF43438; ARG1422.1;
 DR InterPro: IP000635; Viral DNA_bind.
 DR Pfam: PF00747; viral_DNA_bp_1.
 SQ SEQUENCE 1191 AA; -130714 MW; 94A425780FB0711 CRC64;
 Query Match 44-3%; Score 2788; DB 12; Length 1191;
 Best Local Similarity 45.1%; Precl. No. 1.3e-230;
 Matches 542; Conservative 237; Mismatches 395; Indels 28; Gaps 12;
 Qy 1 MENTQKTVVPTPLGTYTACRVEDLLEBISFLARSTSDSLLALPMLRNLTVERKTFPS 60
 Db 1 MDGCKSVRLCGGPICIGIVYATPKCSVPVDELAIAAKSNDCDAVPLVAGITVESDFVW 60
 Y 658 YSSCICPVLSERTLIAVQDMALSQSMMHQQYEANRPTQFQAVIIRRRLDQN 717
 Y 720 GGPFISTRSIITVTLSEGPVSAAPNPTLGQDAPARTFDGDLARVSVEITDIRVKNRVYFSG 779
 b 718 AGFTTSKNITVTLDDQQLISVPDTSKSQHDPLAYMEGLDVKTFEFREFKVKNMFMVG 777
 Y 780 NCTN-LSTAAARAPLVLQASAYQEQKRVDTMLHAGLGEPLKQPHGJLPRGMPPNSKSPNP 838
 b 778 GVSSSTVSDATKSRAAGMELAYOPAKAHVNLQGALKRHTQLPNTVKGNGTPNA 837
 Y 839 QWPTTLLQRNQMPADKLTHEITTAIAKRFTEYYAAINPNLPICGELAQFMANLI 998
 b 838 LWTWILLORNOLPAGILSKDAEINTSFKRFNTSYAMMNVNISPFCFGELAQFVLANI 897
 Y 899 LKYGDSQYLINTLTSITGARRPDPSVLIWIKDVTSAADIEQAKALLEKETENLPE 958
 b 898 LRYCISHKEFFINTISALVATSKERPDAMVLPWIEHLTQGSDVAAPAQOLKVNICDHKE 957
 Y 959 LWTIAFTSTHLYRAAMNOREMVTYGLISISKYHGAGNNRVRQAGNWSGLNGKNTYCPFLT 1018
 b 958 AWZCAFSSTNLVCEIMAKPFVVIAVS1SKHMGAGSTKVFOSGNIGNIMGRNVCIMS 1017
 Y 1019 FDRTRFLIACPRGFICPVTGPSGSNRETTLSDOVRGILIVSGGAIVYOLAUYATTYRAVG 1078
 b 1018 FNRTRFMACPRGFVSSQTGFSQGLKETTVDRAZAVLSDECAGPHAAAYMIALRMVG 1077
 Y 1079 ARAGHMAEDDWLSITDDEFELDLEELHDQIOTLETPTWYEGALEAVKILDEKTTAGDG 1138
 b 1078 DRYQMDLWDMEITNDAYISSLDELNPKQ VEGCEGWSVDAAILAKENYDMSCALPL 1136
 Y 1139 ETPPNLAFNEDSCPSHDTTSNUUNISSNISSTVGLKRPPDEELDLSGPIKHG- 1197
 b 1137 DGPT---FNYDALDENTERRAD----GPSI---LEPTLKPCSD--VFDLEPVFEKRAP 1183
 Y 1198 NITMEM 1203
 b 1184 GLSDVM 1189

RESULT 8

b 986 P0 PRELIMINARY; PRT; 1191 AA.
 D 986 P0 ;
 C 986 P0 ;
 T 01-MAR-2001 (TrEMBLrel. 16, Created)
 T 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 T 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

DE GN MDV042.
 OS Turkey herpesvirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Marek's disease-like viruses.
 NCB_ TaxID=10390;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M65;
 RX MEDLINE=10392152; PubMed=1033706;
 RA Tuinan E. R., Afonso C. L., Lu Z., Zsak L., Rock D.L., Kutish G.F.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
 DR EMBL: AF43438; ARG1422.1;
 DR InterPro: IP000635; Viral DNA_bind.
 DR Pfam: PF00747; viral_DNA_bp_1.
 SQ SEQUENCE 1191 AA; -130714 MW; 94A425780FB0711 CRC64;
 Query Match 44-3%; Score 2788; DB 12; Length 1191;
 Best Local Similarity 45.1%; Precl. No. 1.3e-230;
 Matches 542; Conservative 237; Mismatches 395; Indels 28; Gaps 12;
 Qy 1 MENTQKTVVPTPLGTYTACRVEDLLEBISFLARSTSDSLLALPMLRNLTVERKTFPS 60
 Db 1 MDGCKSVRLCGGPICIGIVYATPKCSVPVDELAIAAKSNDCDAVPLVAGITVESDFVW 60
 Y 61 SLAVSGARTGAGATLKTTSFHSYFSPSVFEGGKHVLESSAAPNLTRACNAARERP 120
 b 61 NVAVAGKETTGLGSGGTTLKVPTTHPVCVYFGCDCIRCTKRPNLTAQDARERF 120
 Y 121 GFSICQGPVDDGAVETTGAEIICTRGLIEPENTILYVTAKEAVFMCNVFLHYGGLD 180
 b 121 GYS-AYSSPAPTAFETTQOQICALENDAQNMVLYLVVTELEKEVYLNCNSLHFQGSDV 179
 Y 181 VHINHGDVTRIPLEPVQLFMPDVNLVDPFENTHRSIGEGGVYPTPYNTGCLUJHDC 240
 b 180 VTNNAVDYRIPYPLHIVLPDNRITNEPFBKRALGEAGIMPKAFYNDSSLCRJHGY 239
 Y 241 VIAPMAYALRVRNNTAVARGAHLAEDENHEGAVLPDITYTQFQSSSGITTAGARRN 300
 b 240 VLSTAVGIVRVRNTDATAARGAHLCPDENHECTTLPDTPFAFTPAAE---TTKQSKM 296
 Y 301 DVNETSKSPSGPERLASIMAAADTALHAEVIFNTGLYEEPTDIKEWPKMFIGMEGTLP 360
 b 297 GKREGSDVS-GGYERTRPASLMSDATAIENVISVYEDIPDYNKWPYICNPVGYAD 355
 Y 361 RLNALGSTYARAVGIGAMVFSNSVYIMTEGASTEGKETSTAPSFRPFQAAFPH 419
 b 356 RIELSATMARIAGLVGAMVFSNSVYIMTEGASTEGKETSTAPSFRPFQAAFPH 415
 Y 420 LAANPOTDRDGHYVL---SOSGGSNTTEPSDYALIICGPAUPALARLFLYLERCDAGA 475
 b 416 LSANPLVDEDGKPVSGENLJSKTSASQEYSDYLLIAGRPOLLARFLYLERCDGA 475
 Y 476 FTGGHG-DALKYTGTGTFSEIPOSICBKHTRVCAPTIVHLRORMPREGQATROPIGVF 534
 b 476 KACHHDLDTVKFSSAIDADMPCELCDKTSRXCANTIKRUYRLPKFGYQCMRGAMGLF 535
 Y 535 GTMSQYQSDCPDIGNAYLIRKPKGDQTEAMKATHQDTRTLFIDLEQERLDRG 594
 b 536 GAMTNINYCDVNAALGSAQFSTLKR--SEGEAARSRSQDQTYLUTVERNMKAKLEGILTC 593
 Y 595 APCESEGUSLIVDPTPRLTRARIEQTTQFWKVLVETRDYKIREELSEATHSMA 654
 b 594 DPTMASADANIRDGNSTRASTMRNIIIESASOLMRNLTIRETMIREGLDANTL 653
 Y 655 LTFDPYSGAFCPTNFYKRTILAVYODLALSQCHCVYGOOGEGRNFRNOFQPVLRRRF 714

Db	654 LATEPYASGICPPVLAFLSRRTIAVQDMAISQCSIWQGQVEARNERTOFOAVLKRRV 713	QY	1 MENTQKTYTVPNGPLGTYAARVEDDLERESFLARSTSDLALLPLRNLTVEKTFPS 60
Qy	715 VDLPNGGJLSTRITVTLSEGPVSAHPNPTGODAPGRTRFDGDLARVSVERDIRVQR 774	Db	1 IMDVGKSYKLVGGPCKYVSHSKDAVEPDSTPAKCTUSEESTIPLTAGLUTSEEFVR 60
Db	714 LEONAGRPTSKTTIVTLEDDQICVPPPSKSOYDSVSNMGEGLVKRTVTEIPRELKVKK 773	QY	61 SLAVVSGARTGAGITLKLTTSHYPSVFPFGKRVPSAANPNTRACHAAREP 120
Qy	775 WPSGNCTN-LSAAARALVGLASAYROEKRVDMLFGFLKOFHGLFPRMMPNS 833	Db	61 NVAAVAGTKTGTLLVNGGTTLKVPTHYHNPVFPYFGDCXKPCSAPNUTKACEYAREP 120
Db	774 VLREGGTTAGAASEBATSKSLAGNEVATORPKTMHVINGPLGAVALKYLFPDVMPNG 833	QY	121 GFRCQGPBPVDAVETGAEICTRGLPEPNTILYLVVTAFLKEAVFMCNFLHYGLDI 180
Qy	834 KSPNPONWFTLICRNOPADKLTHEETTIAAVKRPTEYAYAINFNINLPPCIGELAQFY 893	Db	121 GCPYSAP--GSVENSQKOCOEQLGLDAQTMILUVVTEFKEAFLYLNCSFLYGGSDS 177
Db	834 ATNALAFLWILLRNQNPGATSKSEEDKSFLIKETKSTADMNTNISPICFCGELAQFY 893	QY	181 VHINHGDIVTRIPLPFYVLFMPYVNRVLPDPDPTNTHRSIGCFYVTPPNTGLCHLHD 240
Qy	894 MANLILKYCDHSQYLINTLTSITGARRPDBSVLHWRKDVTSAADIEQAALEKT 953	Db	178 VTINDAEPRRIPIAFLVLPDENVSNEPNSPERPRAIGQAVNBEAFFDNLCKLGY 237
Db	894 LANTILKYCSHKHFINTISALAVASRSPRDPDAVLPWIEPITKQDVAPARQLIASM 953	QY	241 VIAPMAYALRIVRNVTAVARGAHLADEFNEGALEVLPDITYTYF-QSSSGTTTARGR 298
Qy	954 ENIPELWTATPSSTHLYRAAMNQRPMVPLGSIISKTHGAAGNNRVEQAGWSGLNGGRV 1013	Db	238 VLGTDAYGLRVENIDATARGAHLCLDENHEGILLPADTFTTAFTPTADTGKGCGPKACK 297
Db	954 SDFKDINGATPSSTNLVGSIMTCKPFWVIGISISKYHGMAGSTKYEQSGTNAGNIMGGRV 1013	QY	299 RNDVNNTSKPSPGGGPERRLSSIMAADTALEAEVLFNTGUYEETPTDKWPMFTGMEGT 358
Qy	1014 CPMIPTDETRFTIACERGGFLCPVTPGSSENRETTLSDQVRGIVTSGGAMVOAIAYTV 1073	Db	298 RDGAE-----TCGGYTERRTSLSMASDTAIENTIASSTYERPTPDVCKWPICGNPVGY 352
Db	1014 CSIMSPDRTHRYMTCPVGFTEAQPIFSSGKTEITLIDYRMVISEESAPHAIVYMA 1073	QY	359 LPFLNALGSYTAvgAVGIVANEFSNSPALTYEDSGMTPAKDGG-PGESENREYQFAG 417
Db	1074 VRAGAQAQHMAFFDNLTSLTDDEFLARDLLEELHDQIQTLETPTVYEGALEAVTILDEKT 1133	Db	353 TDRVEALSAYCRAVGLGVGANVYSSNSNVVMTVEGEASSSDKESGVAPSPYRFQIAA 412
Qy	1074 LKQGDYRQMLEDWBNTDNEYSLIDELNKQ-VEEAGGGNADAATLAEVMA 1132	QY	418 PHAANPOTDRCGHYLSQS---TGSSNTFSDYDALICFGAPLLARLLFYLERCDA 473
Db	1134 TAGDGETPTNLA芬NDSCEPSSHDTTSVNLNSGSNSGTVPGKRPDEPDLGIP 1193	Db	413 PHAANPOTLVDRGKPSGDAKPLAFSTSSSSISLDLULITAGPCCOLLARFLFYLERCDG 472
2y	1133 MSIPTDGT----FDFDACDENIEGHADGOTISETN-----LKRPP--NMNVDPLEPIP 1179	QY	474 GAFTGGHG-DALKYVGTGTFDEBIPCSLCERTRPVCAHTHTRI-RORMPFFGQAOPIQG 532
Db	1194 IK 1195	Db	473 GANACHEHLDVTKLVSAAIDAMPCCSLCDKTSRVDGKTSRVCYCAHTTVKRLYRIPFGYQURGAMG 532
Qy	1180 EK 1181	QY	533 VFTMNSQYSDCPLGNYAPIYLILAKPGDQEAAATMDQTYRATLERLFDIOEBRLID 592
Db	533 LRCAMTINYCDNAALGSYAQSTLKR--SCEASISVMODTYRATVERVNMALKEGLLA 590	QY	533 LRCAMTINYCDNAALGSYAQSTLKR--SCEASISVMODTYRATVERVNMALKEGLLA 590
Qy	593 RGAPCSSEGGSVIVDDEPTFIRLARIEQTTOPKVKLVETDYKIREGLSEATHS 652	QY	593 RGAPCSSEGGSVIVDDEPTFIRLARIEQTTOPKVKLVETDYKIREGLSEATHS 652
Db	591 CDDPMNNAPAASIKDGPSTHAVSMTRNIEGEASQLMNLTIREVNIREGICDANT 650	QY	653 MALTEDYSGFACPTINFYRTHLAVQDQALSOCHVYGGVEGNFRNOQPVIRR 7.2
Qy	651 LSITIETYSSGFCPUNFLARTRTAVIQMALSOCMSMLMGHQVEARNFRNFTQPVIRR 7.0	Db	651 LSITIETYSSGFCPUNFLARTRTAVIQMALSOCMSMLMGHQVEARNFRNFTQPVIRR 7.0
Db	713 RFDLFGFGFSTRSTRTVTSEGTPSAPNTLGQDAPACRTFEDGLARYTSVETDRYVK 772	QY	713 RYDLOQAGFTSKDQTTALEDQHVAVPDERSQDPTVIMEDGLTRTIEIRELKV 7.0
Db	711 RYDLOQAGFTSKDQTTALEDQHVAVPDERSQDPTVIMEDGLTRTIEIRELKV 7.0	QY	773 NRVFESENCT NLSEARARVNGLASSAYQEKVDMALGALGFLKOPFHGJFPPRGNMP 831
Qy	832 NSKSPMBOWFPTLQLQNONPDKLTHEETTIAVKKPTEEEYAAINFNLPIPICGELAQ 891	Db	771 NRVI FGGGVIGTASDAKSRLANMFBAYOPTKTMVNGPLGPAIKRTHSALFPVNP 830
Db	831 NGAVPNAHWFPTLQLQNONPDKLTHEETTIAVKKPTEEEYAAINFNLPIPICGELAQ 890	QY	832 NSKSPMBOWFPTLQLQNONPDKLTHEETTIAVKKPTEEEYAAINFNLPIPICGELAQ 891
RC	892 FMANILKCYCDHSQVQINTLITSITGARPRDPSVILWIKDVTSAADIEQAKALLE 951	Db	891 FFELANTILKYSKHFFINTISALVAASKRPRDPAIVLPWIDKTIQGDVSTAQQOLIG 950
RC	891 FFELANTILKYSKHFFINTISALVAASKRPRDPAIVLPWIDKTIQGDVSTAQQOLIG 950	QY	952 KBNENLPLWTAFTSHLYRAAMNCRPMVYLGIISKYGAAGNRRVFGQGAWMGLNGGK 1011
RA	951 STADRDAWCATESATINVSSVMSTKPLTIVGVSISKYMGASTKVKFOSGNGNMGGR 1010	Db	951 STADRDAWCATESATINVSSVMSTKPLTIVGVSISKYMGASTKVKFOSGNGNMGGR 1010
RA	952 IYCPPLFEDRPRFLIACPGGFICPVTGESSGRRETTLSDVQGIVIVSGGANYOLAYA 1071	QY	1012 IYCPPLFEDRPRFLIACPGGFICPVTGESSGRRETTLSDVQGIVIVSGGANYOLAYA 1071
RA	952 IYCPPLFEDRPRFLIACPGGFICPVTGESSGRRETTLSDVQGIVIVSGGANYOLAYA 1071	Db	1011 NCISLNSFDRPHRYMACPGFVAEHTGTSIGKEATUMDRARALLSBGGPHAEVYM 1070
RA	1011 NCISLNSFDRPHRYMACPGFVAEHTGTSIGKEATUMDRARALLSBGGPHAEVYM 1070	QY	1072 TVVRAVAGARQHMAFFDWLSLTDEFLARDLSEHHDQIQTLETPTWVGALEAKVILDE 1131

Query Match Score: 2781.5; DB: 12; Length: 1199;
 Best Local Similarity: 44.5%; Pred. NO.: 4.4e-220;
 保守性: 245%; Mismatches: 391; Indels: 35; Gaps: 14;

SEQUENCE FROM N.A.
 STRAIN=FC126;
 RC AFONSO C.L., Tulman E.R., Lu Z., Zaak L., Rock D.L., Kutish G.K.;
 RA Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
 RL J. Virol. 75:971-978 (2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC AFONSO C.L., Tulman E.R., Lu Z., Zaak L., Rock D.L., Kutish G.K.;
 RA Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
 RL AF291866, AAG15671;
 DR InterPro: IPR000635; Viral DNA bind.
 DR Pfam: PF00747; viral_DNA_bp_1;
 SQ SEQUENCE 1190 AA; B450CAF5FF0D928 CRR64;

Db	1071	LALKMVGERVQMELEDWIETGDDVSSLIADLNEQAES-EGGWSIDWAMGIAK---	1125		413	PHIAANPLVDRGKPISGDAPIKLAPESTSSSESYSDYLILACGPCTOLLARLPFLYERCDG 472	
Qy	1132	KTAGDGETPTN-LATNFDSCEPSHHTTSNTLNISGSNIGSTPGLKEPPEDBFLFDLS	1190	Qy	474	GATGGGG-DALKYVGTGFTDEBTPCSCXKTRPVCAHTTIVHRLRORMPRGOATROPIC 532	
Db	1126	ETAASAFIPSNGCPTNEYEENDDGVBEKRS---GHSLSBG---GTRKP---NLISVFDLD	1175	Db	473	GANACHEHDLDVKLVSEASIDADMPCELCDKISRVCAHTTIVKRLPFGYOMRGANG 532	
2y	1191	GIPIKHGN1 1199		Qy	533	VROMNSOYSCDPLCNYAPLILRKPGDQEBAAKATMDOTYRATERLFIDBOERLLD 592	
Db	1176	PVPERAAV 1184		Db	533	LGAMTNNYCVDNALGSAVYQSTLRL--SREASISVMQTYRATERVYKALEKGJIA 590	
	RESULT 10			Qy	593	RGAPCSSEGSSVIVDHPPTFRLDTLRAEQITQFKVLUVTDYKIREGUSSEATHS 652	
ID	Q9E1G3	PRELIMINARY;	PRTR; 1190 AA.	Db	591	CDDPMNMAPADASIKDAGSFTHAVSMRNLTIEGEASOLMNLTECRESTREGLCDANT 650	
IC	Q9E1G3;			Qy	653	MALTFDPYSGFCPCPTNFLYERTHLAVQDQLALSQRCVYQQVEGRFRNQIOPVIRR 712	
JT	01-MAR-2001	(TREMBrel. 16, Created)		Db	651	1SLTIEBYSSESCFPVLSFELLARTTIAVQDMLALSOCMSMLMHGQVEARNFRTOFOSVLLR 710	
JT	01-DEC-2001	(TREMBrel. 16, Last sequence update)		Qy	713	REVDLNGGPFSSTRSRTSVTUELEGPSAPNFTLGODAPASRTFEDDLARYSVEVDIRYK 772	
JE	UL29	major DNA binding protein.		Db	711	RVTELONAGAPTSKQDITATELDQHAYVDPDERSQSYDPTVIMMEDLTVTRTIEIRELYK 770	
JN				Qy	773	MRVVFSGNCT-NLSEARARLVLGASAYORQERKVDMJHGAGLTLKOFHGLLFFRGMP 831	
JC		Meleagrid herpesvirus 1 (herpesvirus of turkeys).		Db	771	MRVIFGGGUGTGASDAKSLSANNLEAYOPTKTMVINGPGLPAKRYHSHLPNVNMP 830	
JC		Viruses; dsDNA viruses, no RNA stage; Herpesviridae;		Qy	832	MSKSPNPQWQWTLLQRNONPADKLTHEETTIAVAKRPEEEYAAINFNPPICGELAQ 891	
JC		Alphaherpesvirinae; Marek's disease-like viruses.		Db	831	NGAVPNRAHWPTLQLRQNQFDLTMQSKEDDKAFLIKKEPTVNAADMNYNISPFGFDIAQ 890	
JN	[1]			Qy	892	RYMANILKCYDHSQYOLINTLTSITGARPRDSSVHLWIRKDVTSAADIEQKAALLE 951	
IN				Db	891	EFLANTILKCYSHKEFFINTISALVAAASKPRDPAIVLPWIDTTITQGDVSLAAQOLIG 950	
IP	SEQUENCE FROM N.A.			Qy	952	KTENIPELWTIAFTSTHLVRAAMNQRPWVGLGSIISKYHGAANRVRQAGHWSGLNGK 1011	
IP	STRAIN=FC126;			Db	951	TIAEDRKDAWATATESATNLVGSWSTKPLIVIGVSIISKYHGMASTKYOFGNTGNGR 1010	
IP	MEDLINE=21195611; PubMed=11297687;			Qy	1012	NVCPLIFTEDTERRFIACPRGGFICPVTGCPSSGENRETTLSDOVGIIVSGAMVQLLAY 1071	
IP	A Kingham B.F., Zelnik V., Kopacek J., Majerciak V., Ney E.,			Db	1011	NVCLSMFSDEHTRTMACEPVGFAEHTGFSGLKEATLMDRARIUSBEGGAPHAEVYM 1070	
IP	Schmidt C.J.,			Qy	1072	TVVRVAGRBRQHMADFDDNLSLTDBEFLARDLETHHDQJQTQLETPWVTEGALEAVKILDE 1131	
IP	"The genome of herpesvirus of turkeys: comparative analysis with			Db	1071	IALKONGBERYRQMLEDWAIBTGDYISSLIADNEAQES-EGGNEIDVANGLAK--- 1125	
IP	Marek's disease viruses";			Qy	1132	KTTADGEPEPTN-LAFENPDSCEPSHDTSVNVNATISGSNISGSGTVPGLKRPPEDEDFLS 1190	
IP	J. Gen. Virol. 82:1123-1135 (2001).			Db	1126	ETAAASAAFTPSNGTENTYEWDDCVKPSD---GHSLSSEG---GHSLSSEG 1175	
IP	EMBL:AF282130; AAC30069_1; -			Qy	1191	GIPIKHGN1 1199	
IP	InferPro; IPRED0635; Viral DNA_bind.			Db	1176	PVPERAAV 1184	
IP	PF00747; viral_DNA_bp; 1.				RESULT 11		
IP	SEQUENCE 1190 AA; 128952 MW;			Q9IB6	Q9IB6 PRELIMINARY;	PPT; 1191 AA.	
				AC	O9IB6		
				DT	O9IB6; 01-OCT-2000 (TREMBrel. 15, Created)		
				DT	01-OCT-2000 (TREMBrel. 15, Last sequence update)		
				DT	01-DEC-2001 (TREMBrel. 19, Last annotation update)		
				DE	UL29.		
				GN	Turkey herpesvirus.		
				OS	Viruses; dsDNA viruses, no RNA stage; Herpesviridae.		
				OC	Alphaherpesvirinae; Marek's disease-like viruses.		
				NCBI TaxID=10390;			
				RN	[1]		
				RP	SEQUENCE FROM N.A.		
				RC			
				RX	STRAIN=G;		
				RA	MEDLINE=92237304; PubMed=1315046;		
				RA	Jones D., Lee L., Liu J.L., Kung H.J., Tillotson J.K.,		

"Marek disease virus encodes a basic-leucine zipper gene resembling the fos/jun oncogenes that is highly expressed in lymphoblastoid tumors."	Qy 768 DIRVAVRVVPSGNCNTN-LISAAARARLVLGLASAYQRQEGRVDMILRGANGFLLOFGRGLFP 826
RT	Db 767 EUKYVKVLEGGCAGAASEATKPLAGKTAQVRPTTMTNGLGAFVBYHTLFP 826
RT	RN [2]
RL	SEQUENCE FROM N.A.
RP	RC STRAIN=GA;
RN	MEIDLINE=20823955; PubMed=10823954;
RC	RA Lee L.F., Sui D., Ren D., Kamil J.J., Kung H.J., Witter R.L.;
RA	RA in the complete unique long sequence of Marek's disease virus.";
RT	RT Organization of the GA strain of Marek's disease virus.";
RL	RL Proc. Natl. Acad. Sci. U.S.A. 97:6091-6096(2000);
EMBL	EMBL: AF147806; AAC66766.1;
DR	DR InterPro: IPR00635; Viral_DNA_bind.
Pfam	DR PF00747; viral_DNA_bp_1;
SEQUENCE	SEQUENCE 119 AA; 130946 MW; AC00A015880079A3 CRC64;
Query Match	Score 43.1%; Score 2710; DB 12; Length 119;
Best Local Similarity	44.1%; Pred. No. 3.6e-214;
Matches	53; Conservative 236; Mismatches 400; Indels 42; Gaps 16;
Qy 1 MENTOKTIVVPTGPGIVYACRVBDLDEEISF -----LAARTSDSLLPLMLNLT 54	Db 1067 1 MDGCKSVKLGGPPI-TMRR-----PVNPAPWNCDLAAKNSCDCATLPTVAGLTIV 53
Qy 55 EKTTTSSLAIVSGARTTGAGATGATLKLTTSHFYPSEVFHGGKHLVLPSSAAPNLTRACN 114	Db 1067 1067 AAVYMLAKWKGDEVRONELEDNAMEITDVESSLIDELNKQ-VEREAGGNNADAANTLA 1125
Qy 54 ESDPWVNAVAVGATPKTTGKGSGTTLKVPVTHAPCIVFVFIGDCIKPTCAPIKTRACD 113	Db 1127 KILDEKTTAGDGEGPTNTAFAFNDSCEPEHDTSNVNLNSGSHTSGSVTPGKLRPPEDEL 1186
Qy 115 AARRERGGFSRCCGPFDGAVETGAICITRGLPENTILYLVVTLALKFKEAVMCNVLH 174	Db 1126 KEMYTMAMSIPTDQPT---FDPDACDENLEGRADGQTSETN----JAKP--NMNV 1172
Qy 114 LARRERFGIS-AIYSSPAPTAFETGQQQICBALENDQAQNVMLYLVVTEFKETYLNCSFLH 172	Db 1187 FDLSGSIPIKHGN 1199
Do 175 YGCFDIVHNHGDUVIRIPVQLEMPDWANLMPDPFNTTHHSIGCEGVYPTPFTNTGIC 234	Db 1173 FDLEPIPEKRAPV 1185
Qy 173 PGGDVVTINNAVDRRIPIYLHVLPDNRTNEPSEKPKPAIGAIMKAFYNDSLC 232	RESULT 12
Db 235 HLHDVCIAPMAYALRVNNTAVARGAHLAEPENHGAVALPPDITYFQSSSSGTTA 294	Q9Q363 ID O9QH63 PRELIMINARY;
Qy 233 RLHGYVISTSTAVGLRVENIDAARGAHLCPENHGTLLIPADTTTAAFP-----TT 289	AC Q9QH63,
Db 295 RGARRNDVNSTSSEPSGPFERLAMIAADPAHLHAEVTFNGTYBPTDJKENMFIG 354	DT 01-MAY-2000 (TREMBLrel. 13; Created)
Qy 290 KGGSXMGKREGSDVS-GGYERYERTASLYMASDATAIENVISASVYEDPIPDVNRKPYICN 348	DT 01-MAY-2000 (TREMBLrel. 13; Last sequence update)
Db 355 MEGTLPLRNALGSYTARYAVGIGAMVSPNSPAIALTLEYEDSGMTEAKD-GGPGPSENRFY 413	DT 01-JUN-2001 (TREMBLrel. 12; Last annotation update)
Qy 349 PVGYADRBEALSAMYARVAGLVAMVSSNSVYMMVEGEAESTETSTTAPSYRFB 408	DB 1 DNA binding protein.
Db 414 QPGPHIAANPOTDRDGHVLSS---SQSTGSSNTTEPSYDYLALICGFAPALARFYYLE 469	OS Gallid herpesvirus 1.
Qy 409 QIAPPHUSANPFLYDGDGXPSGNLNSKTSAAQSYDYLALCFCPCQLARFYYLE 468	OC dsDNA viruses, no RNA stage; Herpesviridae;
Db 470 RCDDGAFTCGHG-DALKYVUTGPFDSIEPCSLCEKETTRPVCAHTTVHLRQMRPQATR 528	OC Alphaherpesvirinae; Infectious laryngotracheitis-like viruses.
Qy 469 RCDDGAHKSHDDTDVKEVSSA1ADMPCELCBKTTSR1CYAHTTICKVRLPKUGYMR 528	OX NCBI_TaxID=10386;
Qy 529 QPGVFGTMNSQYSDCBLGNYAPYLIRKPGDQTEAKATMQDTRATERLFDLQE 583	RN SEQUENCE FROM N.A.
Db 529 GANGLFGAMNTNNDVNAVALVST-AQFSTLKR--SEGEA8RSYMDQTYRLTVERMGALEKE 586	RP STRAIN=Z-A;
Qy 589 RLDDRGAPSSBELLSSVTDHPFRILLDTGARIEOTTQMKVLTETREYNTRGVG 645	RC Johnson M.A.
Db 587 GLJTCDDTNMSADANTRDGNSFIRASTMNTTSEASLQMLNLTETREYNTRGVG 645	RT "Sequence of the infectious laryngotracheitis virus (SA-2 strain) unique long region 1128 to the EN143;"
Qy 648 EATHSMALTFDPSGAFCPITNFVLRKTHLAVYQDLSOCHCVFYGQYEGRVRNFOFO 707	RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases
Db 647 DANHLLSLAVEPAVPSGICPVIAFLSRSQSTVAQGQVBARNRTQO 706	DR EMBL: AF168792; AAD56501; InterPro: IPK00635; Pfn: PH00747; viral_DNA_bp_1; SQ SEQUENCE 999 AA; 108977 MW;
Qy 708 PVLLRRFFVLDLNGGFIISTRSTVTLSSGPVSAAPNPTIGQDAPAGTDFGDLARYSEVER 767	Query Match 23.0%; Score 1445; DB 12; Length 999; Best Local Similarity 30.5%; Pred. No. 8.7e-110; Matches 351; Conservative 177; Mismatches 376; Indels 246; Gaps 18;
Db 707 AVLKKRVLIELQNAGFTSKITVTLDDQOQICVPDPSKSQYDVSISMEGLDVKVTEIFR 766	Qy 13 GPGLGYYACRVEDDLDLBEIISFLAARSTDSD-DIALLPIMRNLTVEKFTFTSSIAVSGARTT 71
Qy 589 RLDDRGAPSSBELLSSVTDHPFRILLDTGARIEOTTQMKVLTETREYNTRGVG 645	Db 22 GPGLGYYVVARPREIDGKSWKLCKASDQPSAIAPIRGILVTESDKPNAAVIGTKSS 81
Qy 648 EATHSMALTFDPSGAFCPITNFVLRKTHLAVYQDLSOCHCVFYGQYEGRVRNFOFO 707	Qy 72 GLAGAGTIIKLTISHEPSVYFHGGKRVLSSAAPPNLTRACNAABREFGSRC-QGPPV 130
Db 647 DANHLLSLAVEPAVPSGICPVIAFLSRSQSTVAQGQVBARNRTQO 706	Db 82 GIVGGNCNTAILSFCHETTVTFHGGCSCIPSFTNLTKCEDAERFGFSSLPPNGPV 141
Qy 708 PVLLRRFFVLDLNGGFIISTRSTVTLSSGPVSAAPNPTIGQDAPAGTDFGDLARYSEVER 767	Qy 131 DGAVETGCAETRIGLEPENTILLYVTAFLKEAVFMCNVFLHYGDLIWHGDVIR 190
Db 707 AVLKKRVLIELQNAGFTSKITVTLDDQOQICVPDPSKSQYDVSISMEGLDVKVTEIFR 766	Db 142 PNLIECTCBEICKSLNDPDCMILYLVVAEEFCBAVVCNTIHYGSVERLYINSKHKVTR 201

191 IPIFPVQLEMPDVN-RLVPDPFNTTHSISIGEGVYVPTFYNTGLCLHTHDCVTAPEVAL 249
 202 CRNPFDNSRNIGEGCIVPTLFLKALNRJHGATGQSL 261
 250 RVRNTAVARGAAHLADEFNHECAVLPPDITYTYFQSSSSGTITGARRNDVNSTSKPS 309
 262 RTRDLEAVARGAAMLADGSFEGCVLTSKTFEL-----AQNIQQTTPPK 307
 310 PSGGFPERLASMAADTALHARPVNTGIVETPTD----IKEYWPMFIGMEGTLPRLN 363
 308 QNVEVERAACSIAEELALATRISVSCAPY--PFDNGCGLPYEWPLFCDKTNERVQ 364
 364 ALGSYPARVAGYTGAKEYFSNPNALYTIVEDSGMTIAKDGCGF--PFSFRVYQFGPH 419
 365 ALDRFSEELAGIVGAIAFSNDSPLYAEVVDGAPF----PGERASSGLTRFYLTLGLH 419
 420 LAANPOTDRDGH---VLSQSQTGSNTTEFSVDYLALICFGCAPLLFLYLERCDAG 474
 420 LGCPDYDSGRVWBTCPSLPTONTFEYDPEHLLAYACGSPSEPLLISRIFTYLECSKQ 479
 475 AFTGGHGDALKVYTGFDSEIPCS----LCEKHTRPVCAHTTVERLQEMPRRGQAT 527
 480 LSTRPN---MNLISSIGNSKAPCTGAPECRWCNDSESRNHCTRYNQRLRTRPRTMR 536
 528 ROPIGFTGMQSODPLNYAPLILRKEDQTAATAKMDTYRATHERLFLFDLEQ 587
 537 RSPIAVIGAVGDDYTIDOLGNFAPSQLRKAEGD-CEPAZAVMDTYRIGIWRVFLYLV 595
 588 ERLLDGAPCSSEGLESSVIVDHPTTERILDLTLRAREOTTQFMKVLYVTDYKREGLS 647
 595 EGLINKDTGDETHNSSL---AEIUTVFERIQNFVSEGSCKPFIASLGTVAYHVBHL 651
 648 EATHSMALTEDPYSGAFCPTINFLYVRTHLAVVQDIALSOCHCVYGOQVEGRNPNQFO 707
 652 SIVHTFG-SPEPYSSAYCPVSLLCQTRSLTFQDILSQIHGTTRQPSKQSSSAM 710
 708 PVLBRREVDLNGGFLSTRSLTIVTLESEGVPISAPNPLQDAPAGTFDGLARVSIEVIR 767
 711 PALRSAFNGMLDKGFSGKYEPFSSVPCNTNAPDTMTRPNTEQATIQAQEYESLSRQGVYKLX 770
 768 DIRVKRNVVFSGNCNULSEAARARLVGLASAYQRQEKRVDMRLGAGFLLKQFHGILFPR 827
 771 EFKVKNRKYVNGSSSNS-----788
 828 GMPPNSKSPNPQWFMTLQLQRNQMPAIDLTHEITIAAVKRFTEEYAAINFNLPPTCIG 887
 789 -----788
 888 ELAQFYMANLILKYCDHSOYLINTLTSITLGARRPRDPSSVILHWRKDVTSAADLTQAK 947
 789 -----788
 948 ALLEKTENLPELWTATFSTTHLVRAMNQRPMVUJGGISISKYHAGAGNNRVEQAGNWSGL 1007
 789 -----CRYVLLSPRIVMGLIMSKYVGQONSTUVFQSNWGSGL 826
 1008 --NGGKRYCPLFTFDTRRFLIACPGGGICPVGPTGSSNRETTLSQVRGIIVSGGM 1064
 827 IGNSSCSQSTSNSALNTDPRVKFALACKRVGTILSSQQVASFOENTIASQVRSLLDAGGS- 885
 1065 VQLAIYATVRAVGRQAHQNAFDMMSLTDEFIARLDEELHDQIOTLTPNTVGEALE 1124
 886 -PTSITAVLRVLGEKGMDITAETWWMAIDDKYLVNLVELREDISGS-QNGWSVATADS 943
 1125 AVKILEDKTT 1134
 944 MLESLEGDT 953
 PRELIMINARY; PRT; 1132 AA.
 ULT 13
 913 040913
 040913;

DT	01-JAN-1998	(TREMBLrel. 05, Created)
DT	01-JAN-1998	(TREMBLrel. 05, Last sequence update)
DT	01-MAR-2003	(TREMBLrel. 23, Last annotation update)
DB	ORF 06.	Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
OS	OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC	NCBI TaxID:37296;	Gammaherpesvirinae; Rhadinovirus.
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=9296220; PubMed=9151804;	
RA	Neipel F., Albrecht J.C., Fleckenstein B.;	
RT	"Cell-homologous genes in the Kaposi's sarcoma-associated rhadinovirus human herpesvirus 8: determinants of its pathogenicity?";	
RT	J. Virol. 71:4187-4192(1997).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RA	Neipel F., Albrecht J.-C., Ensser A., Huang Y.-Q., Li J.J., Friedman-Kien A.E., Fleckenstein B.;	
RT	"The Genome of human herpesvirus 8 cloned from Kaposi's sarcoma.";	
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.	
DR	EMBL: U93872; AA822603.1;	
DR	InterPro: IPR002144; Hbr_SerP-site.	
DR	InterPro: IPR001933; Mitoch_carrier.	
DR	IPR000635; Viral_DNA_bind.	
DR	Pfam: PF00074; viral_DNA_bp_1.	
DR	PS00215; MITOCH_CARRIER; 1.	
DR	PROSITE: PS00589; PTS_HPR_SER; 1.	
SQ	SEQUENCE 1132 AA, T25394 MW, CF88B4F451919CBF CRC64;	
Query Match	10.9%	Score: 603.5 ; DB: 12; Length: 1132;
Best Local Similarity	23.5%	Pred. No. 9e-47;
Matches	301	Conservative 196; Mismatches 500; Indels 285; Gaps
Qy	11 PTGPGGYYACRYVEDLDLLEISFLAARSTDSIDLPLMRNLTVTEKPTFSSLAIVSGART	
Db	18 PTGPGGYYAYLTHNFPGEASLIGNGYPEAKVWPSLPLHIGLTVESDPELNATCAVH----	
Qy	71 TGLAGAGITKLTKTSHFPPSVTFHGGKHVLPPSAAAPNLTTRACNAARERFGFRCQGPV	
Db	74 KIDATTAVSKLTSYH--REAVIYFHNTLHFQPIFQGKGLEKLCRESELFGSTFVBOHQ	
Qy	131 DGAIVETTAEICTRGLPEPNTLYVLTPEAVMCFNVLHYGGIDI-----VH	
Db	132 KGTI--WSPEACPOLPCANE-TFMAVIYTEGGKERL-----YGGKLVPSPQTTPVH	
Qy	183 INHGDVTRIPLPVQVOLPMDVNRVPDDEFTNHRSIGEGFVYVTPPEFTNTGCHLIDHCVI	
Db	181 IGHQATKPLDEDLQPSRQEL-----CRFTNPDIIRYLHDHSIF	
Qy	243 APMAVALVRVNTTAVARGAAHIAFDENINHEGAVLPPDITYTF-QSSSSGTTTARGARRND	
Db	223 TGAQAAQARVKDQDSTVHQ-ASERQF--VHDQXZKIPKLVQAQDPOQASRGTDG-----	
Qy	302 VNSTSKB9PSGGPFERRASIMADTAHLAEEVFTNTG-YEETPTDIKE-----WMPFIG	
Db	272 -----STLMVTD-SLVAELGMSYGJSFIEQPSQDSECEVNTYDTWPFFEN	
Qy	355 MEGTLPLRNALSSYTARVAGVIGAMWESPNSALTYTEVEDSSMTAEKGDPGSFNRFYQ	
Db	314 CTFPDARIALRFLWHAEQALHGQFAANSVILTRV--AKLPQKQNRGDAMNFSYFL	
Qy	415 FAGPHIAANPQDQRDG----HVLSQSTGSENTFEDVYLALICFGAPLARLFYLE	
Db	372 QFLGGLYSEATKENGASAFAKCPVPSALDGSS--VTLQHAYASSSPHILARMCYLQ	
Qy	470 RCDAGAFTGGHD-DALKYVTGTFDSEIPCSUKEKHPRVCAHTTYVHLRQMPR-RGQAT	
Db	429 FLPHHKNTNSQQYNVVDY-GRAAPSMCDLQGQCPAVINTLFRMKDRFPVLSNVK	
Qy	528 ROPIGVFTGMNSQYSDCDPIGNAYPLLRKGDQEAAKTMQDTYRATERLFLDFLQB	
Db	488 RDPPVATGTAQG-TYDLEIILGFATPREFREGNPYEDAP--KTYWYQCONI-----T	

QY	588 ERULDGAPCSSEBGLSSVIVDHPTRILDTLAR1EQTTQFMKVLVTRDKIREGLS 647	Parry J.P., Peruzzi D., Edelman I.S., Chang Y., Moore P.S.,
Db	539 EKLASGISEGGDALRLIVDIPSPVKGIDSTVEAELKFICMKI-NNTYFRENK 597	"Nucleotide sequence of the Kaposi's sarcoma-associated herpesvirus (HHV8)." Proc. Natl. Acad. Sci. U.S.A. 93:14862-14867 (1996).
QY	648 EATHSMALTFDFPSGAFCPITNLFLVERTHLIAVVDOLIALSOCHCIFYGQ----- 695	[3]
Db	598 SVHILQFACNYWQACPVPFLTLYKSLLTVIQDILCTS-CMIEQDNPAVGIVPSEW 655	SEQUENCE FROM N.A.
QY	648 EATHSMALTFDFPSGAFCPITNLFLVERTHLIAVVDOLIALSOCHCIFYGQ----- 695	Russo J.J., Bohenzky R.A., Chien M.-C., Chen J., Yan M., Maddalena D., Parry J.P., Peruzzi D., Edelman I.S., Chang Y., Moore P.S.,
Db	598 SVHILQFACNYWQACPVPFLTLYKSLLTVIQDILCTS-CMIEQDNPAVGIVPSEW 655	Submitted (MAY-1997) to the ENBL/GenBank/DDJB databases.
QY	696 -----OQEGRNFR-----NGFISTRSTIVTUSE 734	EMBL: U75658; IAC57831; -
Db	656 LKQHFQIMWTFNFKGACPDKGATIGGLKIVQSMFCDLFDDAATGMPAPARNQVR--- 712	DR InterPro: IPR003114; HPr-Ser/Thr site.
QY	735 GPySAPAPNTLGODAPAGRTEFGDGLARVSVEVIRDKQEVVSEARARLVG 794	DR InterPro: IPR000635; Mitoch. carrier.
Db	713 -----IAPAMLYPKTKIKRILIS-NSTG-AESIQAQRMK 747	DR InterPro: IPR00147; Viral DNA binding.
QY	795 LASAYQOEKRVDMHLKAGFLLKPHGLLSPRGMPPNSKSPNPWNFTL-QNQMP-A 852	DR Prosite: PS00215; MITOCH CARRIER; 1.
Db	748 PAS-----QRDSY-----IVCGSPYMKFLNALHKTLFP----- 796	DR Positives: PS00589; PTS_HPR SER; 1.
QY	853 DRXTHEETTA-AVERFTEFINLNPCTIGELAOFYMANLILYCDRHQYQINT 911	SEQ Sequence: 1133 AA; T25653 MW; 029ACPAF4C8A5ECB CRC64;
Db	797 PGWSGEHLTELCNYYKASSQAAPEAINVLDVPTDTSVAKIKNSSILACGQTYFATT 856	QY 11 PPGPLGTYAATRCEDDLEEEISFLARSTSIDLALLPMLNLYTEKTFESSSLAV/NSGART 70
QY	912 L-----TSITGARRPRDPSSVLMWIRKDVTSAADIEQAKALLEKTNLPETMTAFT 965	DB 18 PPGPCGTLYATLTHNPIGEAFLPLHGLTVESDFPINRAVH----
Db	857 LSCLSPRTQLVFAEETP-----HVLGPVLSPP-BYTR----- 694	QY 71 TGLAGAGITLGLTTSFSHFSYFSPSFVFFGKHWLPSAAPNLTRACNAARERFGFSRCQGPPV 130
QY	966 STHLVRRAAMNO-----RPMVIGLISISKYHGAANNRVEFOAGN--WSGINGGKNCV 1014	DB 74 KELDADTAAVSUFLSYH-----REAVIFINTNTHFQPIPIRGKGLKELKCBESRELFGFSTFVEQOH 131
Db	895 SVTIVQSTKLKVQSTVNGRLRPTIVLUVKNTGFEAAVKRSVQNVLEDRDPN 1014	QY 131 DGAEVETGAACTRLCEPENTILMLVVTIAFLKEAVFMCNVFLHYGQLDI-----VH 182
QY	1015 P-----LFTEDTRRFLIACPRGGFCIPVTPGSQNRET-TLSDVURGIVS-GCA 1063	DB 132 KC7TL-----WSPECAPOLPCANE-LMMAVIVTEGKFKEL-----YGGKLVLPVPSQTTPVH 180
Db	955 PESVPPFKNNYSSMLERKRVHMTPVDRYKRVGINSFEEAFAVRSVQNVLEDRDPN 1014	QY 183 INGDVIRIPFPPVQLEMPDNRLVYDPPNTPHRSIGEGVYVPPFVNGTGLCHJLHDCVI 242
QY	1064 MVQLAYATVRAVGARAQHNAFDWLSLTDDE--FLARLLEEDHQIQTLET---- 1115	DB 181 IGEHQAKPFLPYDEDLFGPSAAQEL-----
Db	1015 NKRTVYLEVKGIGSS-----ASLTEDVIVYLGYAVLDEVLSLSTGQAGV 1066	QY 243 AMLAVALRVNTAVARGAHLAFTENHEGAVLPEPDITYF-CESSSGTTARGARRND 301
QY	1116 PWTVEGALEAVKILDEKTTAGGETPTNLAFNEDCSEPSHTTISVNLNSGSN---ISG 1171	DB 223 TGAQARVRQDVSITVQ-ASRQF-----VHDQKIKIPLVQAKDFFPQCASRGDG----- 271
Db	1067 PWTAEYGAVIQC-----IDDCE-----LQFYGPEPCLQG 1099	QY 302 VNSTSXPSPSGFERLASMAADTALHAEVIFTGII-YETBTPTDIKE----WPMFIG 354
QY	1172 SYPGKLRKPPEDDELFDLSGIP 1193	DB 272 -----STLVID-SLVAIGMSGLSFLBGPQPSCEVINYDWPIPN 313
Db	1100 QSVV-----EELFPSPGPV 1113	QY 355 MEGTLPRLNACSYTARVAGUTIGAMVPSNALSALTEVEDSGMTEAKDGGPSPNRFQ 414
QY	PRELIMINARY; PRT; 1133 AA.	DB 314 CPTDPDLRALLEVNHAEOQHIGQOLFAANNSVLYTRV--AKLPOKNGQDANNYNSFYL 371
Db	(TRNMBLrel. 03, Created)	QY 415 FAGPHLAANPQTDRG-----HVLSSQTSQSSNTFSDVYLLCFCGAPLLARLFLYE 469
QY	(TRNMBLrel. 03, Last sequence update)	DB 372 ORGLGVLSEATVTKENGASAFFGVPSALDSS-----YTQHLHAYASSFSPLLARMCYLQ 428
Db	(TRNMBLrel. 23, Last annotation update)	QY 470 RCDAAGTGGG-G-DALKYVTCFTDSSHIPSCLCEKTRTRPVCAHTVHLRQMRP-EFGAT 527
DE	ORF 6.	DB 429 FLEPHKONTNSQSYYNVDYV-GTAASQMCDCQGCPAVCINTLYFRMKDRFPVLSVTK 487
OS	Kaposi's sarcoma-associated herpesvirus (KSHV) ("Human herpesvirus 8").	QY 528 RPIGIGTGMMSQYSQCDPQGNYAPVLLRKPGDTEKAATMDTYRATLERLIFIDEQ 587
OC	Virusses: DNA viruses, no RNA stage; Herpesviridae;	DB 488 RDYVIGTGTG-YNDBLILQNFATREREBEGNIVEDAP---KTYWQLCQNT---T 548
OC	Gammaherpesvirinae; Rhadinovirus.	QY 588 ERLLDRGAPCSSEGSSGLSSVIVHPTPRILDTLARIBOTTOQKVLYTRDKIREGUS 647
OX	NCBI TaxID:7296;	DB 539 EXCLASMISEGGDAELFLTIVLIVIPSPVFKVQGIDSVEAELKFICMKI-NNTYFRENK 597
RP	SEQUENCE FROM N.A.	QY 648 BATHSMALTFDPSGAFCPITNLFLVERTHLIAVVDOLIALSOCHCIFYGQ-----
RX	Medline:9704384; PubMed=9939871;	DB 598 SYTHLQFACYVWQACPVPVLTLYKQIYCICLTS-CMYEQDNPAVGIVPSNM 655
RA	Moore P.S., Boshoff C., Weiss R.A., Chang Y.	QY 656 LKHMFQTMWTFNFKGACFDKGATGGBLKTVQSMCDELFTDAAGGMAPPARNQVR--- 712
RT	"Molecular mimicry of human cytokine and cytokine response pathway genes by KSHV";	QY 656 LKHMFQTMWTFNFKGACFDKGATGGBLKTVQSMCDELFTDAAGGMAPPARNQVR--- 712
RT	Science 274:1739-1744 (1996).	QY 656 LKHMFQTMWTFNFKGACFDKGATGGBLKTVQSMCDELFTDAAGGMAPPARNQVR--- 712
RT	SEQUENCE FROM N.A.	QY 656 LKHMFQTMWTFNFKGACFDKGATGGBLKTVQSMCDELFTDAAGGMAPPARNQVR--- 712
RA	Medline:97121480; PubMed=9962146;	QY 656 LKHMFQTMWTFNFKGACFDKGATGGBLKTVQSMCDELFTDAAGGMAPPARNQVR--- 712

735 GPVSNPPTLGDAPAGRTFDGDLARSVVEVIRDIRKVRVVFSGNCINLSEAARPLVG 794
 713 --- :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 Db :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 795 LASAYORQEKRVDMLHGALEGFLKQPHGLFPNMPNPNSPNPQFWTLL-QNQMP-A 852
 748 PAS --- QRSY --- IVGGPKMFLNALHKTLFP --- STKTSALYLHKIGQTKNPIL 796
 853 DKLTIEETTIA-AVERFTEYAAINF1NPIPTCIGELAFQYMLNLYCDISQYLINT 911
 797 PGVSGSHTELNCNYKASSQAPEEIVNDLVDTITSYAKIANSILRACGQTYATT 856
 912 L---TSITGARRPRDPSVHLWIRKDVTSAADIEOAKALLEKTENLPEWTTAFT 965
 857 LSCLSPNTQLUPATEIP---HVLGPGLSSPD-EYRAK---VAGR 894
 966 STHLYRAAMNO---RPMVYLGISISKYHGAGNNRFGAGN---WSGLNGKQVC 1014
 895 SVTIQSTLKOAVSINGRLRPIITPLVNVKTYGTSVFCANLGYFSGRGDVTRL 954
 1015 P---LETFDRTRTIAACBRGGFCICPTCPSSGNRET TISQDVRGJIVSGAM 1064
 955 PESVPEFKONVSSMURKRYHTIMPLVDRILVKRIGINSSEFAEVRSVQVLEEDRD-- 1012
 1065 VQLAIYATVRAVGARAQHMAFDWLSLTDDE-FLARDLLEHDQ1QTLET---P 1116
 1013 -NPNUPKTVYLEL---VKPPRWSQSASLITEDIYFLGAVLDEVLSLSTVQGAVP 1068
 1117 WTEGALLEAVKILLDEKTTAGDEPTLNANFEDCEPSHSDTTSNVLNLSGSN---ISGS 1172
 1069 WTAEGVASV1QDI-----IDOE-----LQFVGPEBPCLIQQ 1101
 1173 TVPGKLKRPPDDDELFDLSGIP 1193
 1102 SVV-----EEFPSPGPV 1114

RESULT 15
 9926 PRELIMINARY; PRT; 1103 AA.
 C Q992Z6;
 T 01-JUN-2001 (TREMBLrel. 17, Created)
 T 01-MAR-2003 (TREMBLrel. 17, Last sequence update)
 E Single-stranded DNA-binding protein.
 6 Murid herpesvirus 4.
 C Viruses; dsDNA viruses; no RNA stage; Herpesviridae;
 C Gammaherpesvirinae.
 X NCBI_TaxID=33708;
 N
 P SEQUENCE FROM N.A.
 C STRAIN=MHV76;
 X MEDLINE=2123509; PubMed=11333912;
 A Macrae A.I.; Dutia B.M.; Malligan S.; Brownstein D.G.; Allen D.J.,
 Misrikawa J.; Davison A.J.; Nash A.A.; Stewart J.P.;
 T "Analysis of a novel strain of murine gammaherpesvirus reveals a
 genomic locus important for acute pathogenesis.";
 J. Virol. 75:5315-5327(2001).
 P SEQUENCE FROM N.A.
 C STRAIN=MHV76;
 A Submitted (Nov 2000) to the EMBL/GenBank/DBJ databases.
 R EMBL; AP24455; AAK16659.1;
 R InterPro; IPR02114; Hhr_SerP site.
 R InterPro; IPR000531; TonB_boxC;
 R InterPro; IPR000615; viral DNA bind.
 R Pfam; PF0074; viral DNA bd1;
 R PROSITE; PS00589; PTS_HPR_SER;
 R PROSITE; PS00430; TONB_DEPENDENT_REC_1;
 R DNA-binding.

FT VARIANT 356 356 S -> G.
 SQ SEQUENCE 1103 AA; 1123272 MW; 05635964BCE27C03 CRC64;
 Query Match 10.4%; Score 652.5; DB 12; Length 1103;
 Best Local Similarity 23.9%; Fred. No. 3.1e-44;
 Matches 292; Conservative 194; Mismatches 535; Indels 199; Gaps 47;
 Qy 1 MENTQKITVTPGTLGLYVYACRVEDDLETSFALARSTDSDLALLPLMRNLTVKTFTS 60
 Db 1 MTTNTVCSQALPQGPAVYIYCLVDFPLEASLLSTNFISAKACILPLVGTIVPGRNF 60
 61 SLAVVSGARTTGLAGAGITURKLTSHFYSVFFGGKHIVLPPSSAAPNLTRACNAARERF 120
 61 NYK---APCQXIMFBSLTVPKPT-KYFNAIVFHSNSNKIPPIFHGRGLQLCPATRIM 114
 Qy 121 QFSRCQCPVPGVAGETTGAECLTRIGLEPNTLYLVTTALPKEARVMCNFLHYGGLDI 180
 Db 115 QFSFSFRDASIGQWRGS---THLPQNPKXIGWVVAESPERLICBLANLSSAPMT 169
 181 WHINHEDVIRIPLPFVOLFMFDVNRLVPDPBNTHRSIGEGFVYPTPFYNTGLCHLHD 240
 Db 170 VTINNKLICFKPLHTRFLSD-OIL-----PFYDENTSRTIMYES 208
 Qy 241 VIAPMAVALRVNTNTAVARSAAHLAFDENHEGAFLPPDITYTYFQSSSSSGTTARGARN 300
 Db 209 YXTNTCAQAVRIRDVGLIEALQERAVTEQYTKALASLKY----- 249
 301 DVNSTSKPSPGGERLSSIMAADTALAEVINTGTYETP-TDIXEWPMFTGMECTL 359
 250 ---PAENCGAGNLEYAVIDAVSELAVSHGLALEAPOBESPQNYQWPMF-SEDAT 303
 360 P-RINALGSYTARYAVGIGAMVFSPSNALSILTEVEDSGMTAEDGGGPSPF-NRFQFA 416
 Db 304 BEGRKALQWMSNANIAHQAQVSANSLYTVIKVGQSA-GRAVGKDDDMFLSRYNQH 362
 Qy 417 G-PHLAANPOTDRCGHVLSQSSTGSS---NTEFSVDYLALICFGAPLARLFYLERCD 472
 Db 363 GLGHL-NVORDENNTVYFGVDSCKNGSQTVHHHLAYAAMSMPHLARLCTYLCQMCQ 420
 Qy 473 AG-ATFGGIGDAKXYTGTTEDSB1PCSLCBEKTRPVCAHTTFLRIFORMPR FGQATRQ 529
 Db 421 NQRLASTPQSTTVPAYVTTANTSM-CDCIGGDLPAVCLHTFLRDRFPVTLATQRD 479
 Qy 530 FIGVGTGTMNSQYSDCDPQNYAPPLRKPGDQTEAAKATMDTYTRATLERRFIDLEQER 559
 Db 480 FYYVWT-LAGPHMELLNGFNSP--KEKDDQ-----ENGQRSYWMQNLNTER 527
 Qy 590 LLDRGAPCSSEGLSSVYHPTRRDLTRARIEOTTQFMKVLYVTR-DYKIREGLSE 648
 Db 528 LEEIGVTAATPA-CQLTDVDSLTVFKKIDAVDEETIRFDGLVTRNNINK--ESVKS 584
 Qy 649 ATHSMALTPDPYSGAFCPCITNFVYKRTHLAVYDLSQG-----HCVFYQQEVEGRNF 702
 Db 585 IAHV1QLOQNTNWLPCCP1HNFTRSPETLDDMSFLCLVYTHPENPLNYG--TVSTM 642
 Qy 703 RNOQFQVLRERRFVLDENGFISTRSITYLSEGPVSAARNPTIQDABAGRTFDGLARVS 762
 Db 643 SNHFONLWSNFSK1FDCOLTCSDMRVHTE-TSSEMDNAALEPKRCAI-RUSRAQ 699
 Qy 763 VEIRDIVKRYVVFSGNCINLSEAARARLYGLASAVQERKVRDMHLHAGLGFLLKOPHG 822
 Db 700 IFAPAKLK1K1RIFNSSS--TESTSASFVTKGKAEN-----PITAGPYMHFLQTHK 51
 Qy 823 LLPPRGMPNSKSBNPQNMFW-TLQRNOMP-----ADKLTHEETITAAVYKRFTEYAA 875
 Db 752 QLFP-----TTKTKGAYTIWTEMSSKLQPGSTNKKL EFSYLMTSSLAHDE--- 801
 Qy 876 INFINLPPTCIGELAQFYMANLILKYCDHSQYINTLTSI-----ITGARR 921
 Db 802 VNVLDYVPTPLMAYAKOBLNNAILRLCQTYATTINFLQDTILSMSPLDYPHYAG--- 858
 Qy 922 PRDSSVLLHWIRKDVTSAADIEQAKALLEKENLPLWTAFTSTHLVRAAMNQRPMTV 981

b	859	---PVIV----QPLTYNDMTRDQSATVISSCRED-----PAAVCKNRPIVT	899
y	992	LGISISKYFGAAGNNRYFOAGNNSGLNGRNVCPLFFFDR-----TRR 1024	
b	900	: : : : : : : : : : : : : : : : :	
b	900	IPKMKVYKSGNSNNSIQSGN-MGYPMGRGV-----DRKULPDLPRMRKHNTMSRR 952	
y	1025	FITACPRGGFICPTGPSSGNRBTTLSDQVR----GIVISGAMVQIAVATVRAVGAR 1080	
y	953	: : : : : : : : : : : : : : : : :	
c	953	YAFASPMTESL-LTQPKEKGTVRTYOLEKTRKIHDDITIASNAEPP-NLVIELVRLGHE 1009	
y	1081	AQMHAFFDWLSLTDDPEFLARDLBELHDIIQLETPTWVEANL--BAVKILDEKTAJDG 1138	
c	1010	CATLTSDF----DVEFYLGEGFSMMVNDVMVELETTIKASGCCAFSEESSVQALLPISSPGTS 1063	
y	1139	EPTNLNAPNFDSSEPSHDTT 1158	
c	1064	ESPVETVNITISLPEISVT 1083	

search completed: January 30, 2004, 13:14:03
db time : 52 secs

Qy	181	VHINHGDRVTRIPPEVQLPFMPDNRLVDPFNTTHRSGEGFVYPTPFYNTGLCHLJHDC 240	DnB1_HSVB	STANDARD;	PRT;	1209 AA.
Dn	181	VHINHGDRVTRIPPEVQLPFMPDNRLVDPFNTHRSIGEVYPTPFYNTGLCHLJHDC 240	AC	P26532;		
Qy	241	VIAFMAYALVRVNTTAVARGAAHLAFDENHEGAVLPPDITTYFQSSSGTTARGARRN 300	DT	01-DEC-1992 (Rel. 24, Created)		
Dn	241	VIAFMAYALVRVNTTAVARGAAHLAFDENHEGAVLPPDITTYFQSSSGTTARGARRN 300	DT	01-DEC-1992 (Rel. 24, Last sequence update)		
Dn	241	VIAFMAYALVRVNTTAVARGAAHLAFDENHEGAVLPPDITTYFQSSSGTTARGARRN 300	DT	16-OCT-2001 (Rel. 40, Last annotation update)		
Qy	301	DVNSTSKEPSPGCFERRLASTMADTAIHAEVENTGTYEETPDKEWPMFNGEPLITP 360	DB	Major DNA-binding protein.		
Dn	301	DVNSTSKEPSPGCFERRLASTMADTAIHAEVENTGTYEETPDKEWPMFNGEPLITP 360	GN			
Qy	361	RNALGSTARVAVGIVAMVESSNSALYLTEVEDSGMTAEKGPPSPRFRYQTAGPHL 420	OS	Equine herpesvirus type 1 (strain Ab4P) (EBV-1).		
Dn	361	RNALGSTARVAVGIVAMVESSNSALYLTEVEDSGMTAEKGPPSPRFRYQTAGPHL 420	OC	Viruses; dsDNA viruses; no RNA stage; Herpesviridae;		
Qy	361	RNALGSTARVAVGIVAMVESSNSALYLTEVEDSGMTAEKGPPSPRFRYQTAGPHL 420	OC	Alphaherpesvirinae; Varicellovirus.		
Dn	361	RNALGSTARVAVGIVAMVESSNSALYLTEVEDSGMTAEKGPPSPRFRYQTAGPHL 420	OX	NCBI_TaxID=51520;		
Dn	361	RNALGSTARVAVGIVAMVESSNSALYLTEVEDSGMTAEKGPPSPRFRYQTAGPHL 420	RN	[1]		
Qy	421	ANPQTDRDGAVLSSQSTGSNTTEFSYDYLALLCGFAPIALARLLYLERDAGAFTGGH 480	RX	SEQUENCE FROM N.A. MEDLINE=9229566; Published=1318606;		
Dn	421	ANPQTDRDGAVLSSQSTGSNTTEFSYDYLALLCGFAPIALARLLYLERDAGAFTGGH 480	RA	Telford E.A.R., Watson M.S., McBride K., Davison A.J.;		
Dn	421	ANPQTDRDGAVLSSQSTGSNTTEFSYDYLALLCGFAPIALARLLYLERDAGAFTGGH 480	RT	"The DNA sequence of equine herpesvirus 1."		
Qy	481	GDLKYYTGTGFDEIPCGSLCEKTRPVCAHTTVHLRQRMRPRFGTQTRPQIGVIGTMNSQ 540	RL	Virology 189:304-316 (1992).		
Dn	481	GDLKYYTGTGFDEIPCGSLCEKTRPVCAHTTVHLRQRMRPRFGTQTRPQIGVIGTMNSQ 540	CC	-!- FUNCTION: SINGLE-STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA		
Qy	541	YSPCDPLGNYAPLILRKPGDOTEPAKATMDTYRATLERLFLIDQEERLDRAPCSSE 600	CC	-!- SUBCELLULAR LOCATION: Nuclear (Probable).		
Dn	541	YSPCDPLGNYAPLILRKPGDOTEPAKATMDTYRATLERLFLIDQEERLDRAPCSSE 600	CC	-!- SIMILARITY: BELONGS TO THE HERPESVIRUSES DNA-BINDING PROTEIN FAMILY.		
Qy	601	GLSSVIVDHPTRRILDRLARIBOTTOEMKVLVETRDYKIRELSEATHSMALTFDYY 660	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EBI, outstation of the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
Dn	601	GLSSVIVDHPTRRILDRLARIBOTTOEMKVLVETRDYKIRELSEATHSMALTFDYY 660	CC			
Dn	601	GLSSVIVDHPTRRILDRLARIBOTTOEMKVLVETRDYKIRELSEATHSMALTFDYY 660	CC			
Dn	661	SQACPCTNLFYKRTTHLAVQDIALAQCHCIFYGOVEGANFRNQFQPVIRRFDLFG 720	CC			
Dn	661	SQACPCTNLFYKRTTHLAVQDIALAQCHCIFYGOVEGANFRNQFQPVIRRFDLFG 720	DR	ENBL: M8664; DAB02466.1; -		
Dn	661	SQACPCTNLFYKRTTHLAVQDIALAQCHCIFYGOVEGANFRNQFQPVIRRFDLFG 720	DR	DR: E31698; DNREC4.		
Dn	721	GPI1S1TSITVLSGPYSAPNPTLGODAAGARTFEDDLAISVSYTVDKRVQNSVFFSN 780	DR	InterPro: IPR000635; Viral DNA bind.		
Dn	721	GPI1S1TSITVLSGPYSAPNPTLGODAAGARTFEDDLAISVSYTVDKRVQNSVFFSN 780	DR	PFAM0147: viral DNA bp_1		
Dn	721	GPI1S1TSITVLSGPYSAPNPTLGODAAGARTFEDDLAISVSYTVDKRVQNSVFFSN 780	KW	DNA-binding; DNA replication; Zinc-finger; Nuclear protein.		
Dn	721	GPI1S1TSITVLSGPYSAPNPTLGODAAGARTFEDDLAISVSYTVDKRVQNSVFFSN 780	FT	ZN FING 503 516 C2RC-TYPE.		
Dn	781	CTNLSEARARIVGLASAYORQEKRVDMLHICALGFLKQFHGLLPRGMMPNSNSPNTOW 840	SO	SEQUENCE: 1209 AA; 129982 NM; 1A728PB04484FE95 CRC64;		
Dn	781	CTNLSEARARIVGLASAYORQEKRVDMLHICALGFLKQFHGLLPRGMMPNSNSPNTOW 840	QY	Query Match 57.5%; Score 3620.5; DB: 1; Length 1209;		
Dn	841	FWTLLQRNOMPDKLTTHEETTIAAVKRFEEYAAINFNLPPCTIGELAQFYANLTIK 900	QY	Best Local Similarity 57.1%; Fred. No. 4.7e-270; Matches 698; Conservative 191; Mismatches 299; Indels 35; Gaps 14;		
Dn	841	FWTLLQRNOMPDKLTTHEETTIAAVKRFEEYAAINFNLPPCTIGELAQFYANLTIK 900	DB	1. MENTQTVIVTGTPLGKYYTAKRVEDLDLEISFARARSDSDAIIPLMRNLTEKTTS 60		
Dn	901	YCDHSQYLNPLNTSITTGARRPRDSSVLFMIRDVTSAADIEQAKALLETKENLPNW 960	QY	1. MESAPKTVLSLVPSPLYGIVYARQKASLQGTVSLAARS5D2AVLPVTRGIVTEQFTT 60		
Dn	901	YCDHSQYLNPLNTSITTGARRPRDSSVLFMIRDVTSAADIEQAKALLETKENLPNW 960	DB	2. SLAVISGARTGLAGAGITKLTKLTSFHFYPSVYFHGGKRYLSSAAPNITRACNAAPRF 120		
Dn	961	TTAFTSTHLYRAAMNORPMVTLGISTSKYHGAAGNNRVIQAGNSGLNGKNOVPLFTD 1020	QY	6. IVAVVAGSKTFLG3GJLITKLTPRHFNPAFVYGGSTGASSNAPNITRACAAHRRF 120		
Dn	961	TTAFTSTHLYRAAMNORPMVTLGISTSKYHGAAGNNRVIQAGNSGLNGKNOVPLFTD 1020	DB	12. QFSRCGQPPVYDGAETTGAICTRGLBPENTLYLVTALIACEAVPNCVNPHYGGLDI 180		
Dn	1021	RTRPFTIACPQQFICPVTGPSSGGRFTLSDQRTIVTGGANVQLAYATVRAYCAR 1080	QY	12. VTFIDGDAMKIPVYOLVQYDPMVNRLASPEPNAXHRSGDEFYTSRPENSDICRLLHG 240		
Dn	1021	RTRPFTIACPQQFICPVTGPSSGGRFTLSDQRTIVTGGANVQLAYATVRAYCAR 1080	DB	12. VTFIDGDAMKIPVYOLVQYDPMVNRLASPEPNAXHRSGDEFYTSRPENSDICRLLHG 240		
Dn	1081	AQHMAFDWDLSLTDDEFLARDELLHDQIOTLLEPWTYEGALPKILDEKTAGDSET 1140	QY	241. VMAPMAYALVRVNTAVARGAHLAFLDENHEGSJLPQDVTPLFDSQGNRAGKSGRAQR 300		
Dn	1081	AQHMAFDWDLSLTDDEFLARDELLHDQIOTLLEPWTYEGALPKILDEKTAGDSET 1140	DB	241. VLGPARVALVRVNDVARGAHLAFLDENHEGSJLPQDVTPLFDSQGNRAGKSGRAQR 300		
Dn	1141	PTNLAFNFDSCEPNSHDTSNLYNIGSNISGSTVGLKRPPDEDBLFDLSGITKHGNIT 1200	QY	299. RNDVASTSKESPSGCFERLALSINAADTAHLAFTFNTGIYETPTDKENPMFIGNEGT 3-8		
Dn	1141	PTNLAFNFDSCEPNSHDTSNLYNIGSNISGSTVGLKRPPDEDBLFDLSGITKHGNIT 1200	DB	301. CGD-CGSGSKNSASSSSTERLASYWADLILSDSIMGAGIDBPLPSVBDWPLSSGDDT 359		
Dn	1201	MBM 1203	QY	359. LPRNVALGYSTARYGVIGAMYFPPNSAUYLTYEDVSAMTEANDGGRPSFENFYOAGP 418		
Dn	1201	MBM 1203	DB	360. -BSLELGA-FARISGLYAMMFANSVLYMTEDDGCPADGDGS-NPSYRFYLAAP 417		

RESULT 3		SEQUENCE FROM N.A.	
JBI-HSV1K	DNB1-HSV1K	STANDARD:	PRT: 1196 AA.
P17410;			
01-AUG-1990 (Rel. 15, Created)			
01-AUG-2001 (Rel. 15, Last sequence update)			
16-OCT-2001 (Rel. 40, Last annotation update)			
Major DNA-binding protein (Infected cell protein 8) (ICP 8 protein).			
DBP OR UL29 OR ICP8.			
Herpes simplex virus (type 1 / strain KOS).			
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
Alphaherpesvirinae; Simplexvirus.			
[1] _TaxID=10306;			

530 RGAIGVFTGMNSWYSDCVLGNAYAFAFSALKR - ADGSETARTIMQETYRAATERVMAELET 588
 588 ERLLDRGACPSSBELLSSVTVDDHPTERRLDLTARIQTTTQFMKLVETDVKREGLS 647
 QY :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 Db 589 LQYDQAVETAMSRLELTITNRALHTYNNTRQVDREREVQLMNLVEGINKFKFDGLG 648
 QY :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 Db 648 EATHSMALTEDPSGAFCPINLVLVERTHLAVYQDLALSQCHCVCVNGQQVEGRNFRNQFQ 707
 QY :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 Db 649 EAHAMSSTLDPACGPCLLQLLGRRNLAYQDLALSQCHGFGAGOSVEGRNFRNQFQ 708
 QY :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 708 PVDERRFLDENGIFISTRITVTLSEG-PVSAPNPTLQGDAPGRTFDGDLARYSVEVI 766
 QY :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 Db 709 PVYRRYNDMENGGELSAKLTVALSEGAICPSPSLTAGQTAPEESFEGDVARYTLGFP 768
 QY :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 767 RDYRKVQVFVSGNCNTNSEAARLYGLASAYQREKRVDMLHEGALGFLLKQFGGLLFP 826
 QY :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 Db 769 KERVKSRVLFAGASANSEAARVYLQSAWQKPPKRVDLLGFLGKFLQPKRRAIIP 828
 QY :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 827 RGMPPNSKSPNPWFWTILQRNCMPADKLTHEITIAAVKRFTEYYAAANFINLPPTCI 886
 QY :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 Db 829 NGKPGSNQNPAPNPFWTAQLRQLPAULSRDDETAIFKFKSLDYGAINFINLPAPNNY 888
 QY :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 887 GEHAQFYMANLKYCDHSQYINTTSITCARRPRDPSVLYHWRKDVTSAADIETOA 946
 QY :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 Db 889 SEAMYMTMQNQTYCDHSTYNTNTTATAGSRRPSVQAAAIV--SAQGAGLEAGA 945
 QY :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 947 KRILEKTEENLPWLTWTAFTSTHVRAMMNQRPVMVIGISTSKYHGAGNNRVRQAGNWG 1016
 QY :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 Db 946 RALMDAIDAHPGAWTMSWAFASCNLIRPYMAAPPMWVGLSISKYGMAGNDRVFQGNMAS 1005
 1007 LNGGKVNQVPLTFDRTRFTRFIACPRGFFICPVTVGPGSGNRETTEDQVRLTIVGGAMYQ 1066
 1006 LMGKGNACPLLIIDRTRAKFVLAICPRAFGVCAASLUGGAHESLQLRLISGGAAVA 1065
 2y 1067 IAIYATVYRAVGRAGOMAFDWLSTTDDEPLARDLDEELHDQITQLETPWTVGAL--- 1123
 Db 1066 SSVFVATVKSLSQPTQOLQEWLALDEDTSEENMEETPARLQRNGEWSTDALEYA 1125
 2y 1124 -BAVKILDEKUTTAGDGETPTNLAFEND- -SCEPBDHTTSNVLN1SGSNISGSGTSPGLKRP 1180
 Db 1126 HEAELYLSQNGAGB ----- VFNFGDFGCE ----- DDNATPFGCPGAPAGRCRA 1173
 2y 1181 PEDDELFDLSGTPIKHGNITMEM 1203
 Db 1174 FHGDDPFG-EGPDKKGDLTLM 1195

RESULT 4
 [1] NBI_HSV1 STANDARD; PRT; 1196 AA.
 ID -DNBI_HSV1
 AC P04296;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Major DNA-binding Protein 8 (Infected cell protein 8) (ICP 8 protein).
 DS Herpes simplex virus (type 1 / strain 17).
 DC Alphaherpesvirinae; Simplexvirus.
 XN NCBITaxId=10239;

[2] SEQUENCE FROM N.A.
 UP MEDLINE=88274327; PubMed=2839594;
 IX McGeech D.J., Dairymple M.A., Davison A.J., , Frame M.C., McNab D., Perry L.J., Scott J.E., Taylor P., "The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1.", J. Gen. Virol. 69:1531-1574 (1988).
 IN SEQUENCE FROM N.A.
 UP MEDLINE=8606723; PubMed=2999714;
 IX Quinn J.P., McGeech D.J., "DNA sequence of the region in the genome of herpes simplex virus

RT type 1 containing the genes for DNA polymerase and the major DNA binding Protein.", Nucleic Acids Res. 13:8143-8163 (1985).
 RT
 RL
 RN [3]
 RP SEQUENCE OP 1062-1196 FROM N.A.
 RX MEDLINE=1830232; PubMed=245727;
 RA Hammer Schmidt W., Conraths F., Mankertz J., Buhk H.-J., Pauli G., Ludwig H., "Common epitopes of glycoprotein B map within the major DNA-binding Proteins of bovine herpesvirus type 2 (BHV-2) and herpes simplex virus type 1 (HSV-1).", Virology 165:406-41 (1988).
 RT
 RL
 CC REPLICATION: SINGLE-STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA REPLICATION.
 CC -!
 CC -! SUBCELLULAR LOCATION: NUCLEAR. IN THE ABSENCE OF DNA REPLICATION, FOUND IN THE NUCLEAR FRAMEWORK-ASSOCIATED STRUCTURES (PREPLICATIVE SITES); AS VIRAL DNA REPLICATION PROCEEDS, IT MIGRATES TO GLOBULAR INTRANUCLEAR STRUCTURES (REPLICATION COMPARTMENTS).
 CC -!
 CC -! SIMILARITY: BELONGS TO THE HERPESVIRUSES DNA-BINDING PROTEIN FAMILY.
 CC -!
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.ebi.ac.uk/announce/> or send an email to license@ebi.ac.uk).
 CC -!
 DR X14112; CAA3232; 1;
 DR X03184; CAA2690; 1;
 DR M21631; AAA4577; 1;
 DR P1R: A03790; DIREV1;
 DR InterPro: IPR00635; Viral_DNA_bind.
 DR PE0047; viral_DNA_bp; 1;
 DR DNA-binding; Zinc-finger; Zinc-finger; Nuclear protein.
 DR ZN_FING: 499 512 C2HC-TYPE.
 DR SEQUENCE 1196 AA; 128-49 MW; 453799-62E5B9939 CRC64;
 DR Query Match 49 4%; Score: 3049; DB 1; Length: 1196;
 DR Best Local Similarity 49 8%; Fred: No. 4 -4e-226;
 DR Matches 609; Conservative 199; N-matches 367; Indels 48; Gaps 17;
 DR PT SQ
 QY 1. MENTQKTVT--VPTGPIGTY--ACRVEDLDBTSFLAARSSTDSDAILPLMRNLTVE 55
 DR 1. METKEPKATTKVPPPLGVWYARACPSGIEL -LALISARSSDSDVAVPLVGLTV 58
 DR 56 KIFTSSLAIVVGARTTGLAGAGITKLTKLTHFYSIVFPHGKHVLPSSAAPNLTRACNA 115
 DR 59 SGPEANVAVVGSRTGCPPDGAETTGABCTIGLEPENTILYVLTALFEAVFCNVEFLHY 1175
 DR 116 AREREFPSRCGCPPDGAETTGABCTIGLEPENTILYVLTALFEAVFCNVEFLHY 1175
 DR 119 ARRHFGESTDYPRGDQLKHTGTBALCERGLDDRALLYVTEGEAVCINTFLHL 178
 DR 176 GQDITHINHEDVTRIPLEPVQLEMPDYNPLVPPPTNTHRISGEFYPTPPYNTGLCH 235
 DR 179 GGSDKVIGGEAEVIRIPVFLQIMPDEFVRLPEPFPNRPINR 238
 DR 236 LIHDCTVAPMVAIRVNTAVAGAAHIAFDENHEGAVLPDTITYFQSSSSGTTAR 295
 DR 239 ILFEEATVGPAPAVLRCRNTDAVAAAHIAFDENHEGALPADTFTFAEASQG -KTPR 296
 DR 296 GARRDNVNSTSPSPTGGFERLASTMADTLAHEVENTGIVYEETTDICCEWMFEGM 355
 DR 297 GGR --- -DGCKGKAGGGEORLASMAGDAALESIVSMAYFDEPTDLSAMPFLFGQ 351
 DR 356 EGTLPRILNALGSYTARVAGTIGAMFSPNSAUYTEVEDSGTEAKDQGPGBSFNRYEQF 415
 DR 352 DTAARANA/GAYLRAAGLWGANFVSTSALSHTEVDAGFADPDKSK-SFYRFLV 410
 QY 416 PGHHLAANPOTDRGCHVL --- SSQTGSSNTTEFSYDYLALICRGAGAPIALARLFLY 468

411 PGTHVAAANQVDRDGHVVPGEGRPTAPLVGTO-Q-EFAGEHLLMCGFSPLLAKRMFLFYL 469
 469 ERCDGAFTGCHG-DALKYTGTTSEPSLCKEKTRECAHTVHLRQMRPREGAT 527
 470 ERCDGAIVTGRQENDVFRVADPSQTDVPLNLCFDTAACVTCMLRHPKFASAA 529
 528 RQPQVFGTMNSQSDCDPGLNYAPYLIRKPGQTEAAKATMQCTRATERBLFILEQ 587
 530 RGAGAVFGTMNSMDSDCDVIGNYAFAFSALKR-ARGSETARTINQETVAAERMALEI 588
 588 ERLLDRGAPSSSEGSSV1VDPFPTRILDLRRIEQTTOMKVLVETRDYKIREGLS 647
 589 LQYDQAVPTAMGRERLTINRREALHTVNNYVQVYDREVEQLMNLVEGRNFKFRDGL 648
 Y 767 RDIRYKRTVVFSGCNTLNSEAAARARLVLGASAYQROEKVDMLGHAIGELIKQFGLLPP 826
 b 769 KEJRVKSRLVLAGSANASEAAARLVALQSAYQKPDGRDILQGPGLFELKQFRAIIP 828
 Y 827 RGNVPSNSKSPNPQWETLQLQRNQPADLTHEITTTAAVKRPTEEAANFINLPTCI 886
 829 NGKPGSNOPNPQWETLQLQRNQPADLTHEITTTAAVKRPTEEAANFINLPTCI 888
 887 GEJZQFMANLILKYMCDHSQYLINTLTSIITGARRPRDPSSYLHWIRKDVTSAADIETOA 946
 b 889 SEIAAMYNNQNLRYCQDSTYFNTTAAITASRPRPSVQAANW-SMAGGAEAGA 945
 Y 947 KALLEKTNLPPEWTTAFTSHVRAAMNQRPVWGLGTSISKYHGAGNNRYFOAGNWSG 1006
 b 946 PALMDAVIDAHPGMTSMFASCNLRLPMAARPVVLLSISKYGAGNDEVQAGNWAS 1005
 1007 LNGCKNVNCPLETFDTREFLIACTPRGGFICPVTGPGSSGNRETTLSQDQRETLVSGGMVQ 1066
 Y 1008 LMGGNAAPLTLFDRTRAFVLAICPRAFGVCAASSLGGAHESSLCHOLRGHISEGAAVA 1065
 1067 LATYATVYRAVGARAQEMAFFDMLSLTDDPEFLARDLELHQIIQTLTETPMVGBAL--- 1123
 3 1066 SSFVEVATPKSLGRTRQQEMLVALLDEUYSEEMELTARALRNGNWSTDVALEYA 1125
 1124 -EAVKILDEKTTAGDTGEGTPNLAFNFD--SCBPSHDITSVNIUNISSNISGSTVGLKP 1180
 f 1126 HEAELYSQQLNAGE-----VNFQDFGCE----DDNATPFQGPAGFPAGRKRA 1173
 / 1181 PEDDELPLLSGIPKHNNTMM 1203
 > 1174 FHGDDPFEG-EGFPDKKGDLTDM 1195

RESULT 5
JBI_HSV1F
 DNB_HSV1F STANDARD; PRT; 1196 AA.
 P17469; 01-AUG-1990 (Rel. 15, Created)
 01-AUG-1990 (Rel. 15, Last sequence update)
 16-OCT-2001 (Rel. 40, Last annotation update)
 Major DNA-binding protein (Infected cell protein 8) (ICP 8 protein).
 DNP OR UL29 OR ICP8.
 Herpes simplex virus (type 1 / strain F).
 Viruses: dsDNA viruses, no RNA stage; Herpesviridae;
 NCBII_TaxID10304;
 SEQUENCE FROM N.A.; PubMed=2641793;
 MEDLINE=8806221; Conraths F., Mankertz J., Pauli G., Ludwig H.,
 Hammarschmidt W., Hammerschmidt W., Conraths F., Mankertz J., Pauli G., Ludwig H.,
 P17469; 01-AUG-1990 (Rel. 15, Last sequence update)

RA RT Conservation of a gene cluster including glycoprotein B in bovine herpesvirus type 2 (BoHV 2) and herpes simplex virus type 1 (HSV-1).;
 RT RT
 PT Virology 165:38-40 (1988).
 RL -!
 CC FUNCTION: SINGLE-STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA REPLICATION.
 CC -!
 CC SUBCELLULAR LOCATION: NUCLEAR. IN THE ABSENCE OF DNA REPLICATION, FOUND IN THE NUCLEAR FRAMEWORK-ASSOCIATED STRUCTURES (PREREPPLICATIVE SITES); AS VTPAL DNA REPLICATION PROCEEDS, IT MIGRATES TO GLOBULAR INTRANUCLEAR STRUCTURES (REPLICATION COMPARTMENTS).
 CC -!
 CC SIMILARITY: BELONGS TO THE HERPESVIRUSES DNA-BINDING PROTEIN FAMILY.
 CC DR DR PIR: D29212; DNBNBEH; Inter-Pro: IPR00635; Viral DNA-bind.
 CC DR DR Pfam: PF00747; viral DNA_bp_1; Zinc-finger; Nuclear protein.
 CC KW DNA-binding; DNA replication; Zinc-finger; Nuclear protein.
 CC FT 2C1C-TYPE.
 CC SQ Sequence 1196 AA; 128373 MW; BC82584DDB1C8B2 CRC64;

Query Match Score 48.4%; Score 3047; DB 14; Length 1196;
 Best Local Similarity 49.8%; Pred. No. 6.3e-226;
 Matches 609; Conservative 199; Mismatches 367; Indels 48; Gaps 17;

Qy 1 MENTOKTIVT--VPTGPIGYVV--ACRVEDLILBEISPLAARTSDSDSPLALLMLRNLTVE 55
 Db 1 METKPTATTIVTKVPGPPLVYVVARACPSEGIEI-LAIIASGDSPVAVALVVGITVE 58

Qy 56 KTFETSSLAVYSGARTTGLAGAG-TKLTTSHFYPSPVYFHGGKHLVPLPSSAAAPNLTRACNA 115
 Db 59 SGFEPANVATVVGSSSTTGLGTAVSXLUTPSHVSSTVYFHGGRHLDESTOPNLTRCLER 118

Qy 116 AREFGFSERCQGPVVDGAETVGAIECTRLGLPENTILYLVVTAFLKEAVMCNTVLYH 175
 Db 119 ARRIFGFSDTYTPGDLKHETGEALCERLGDPRDRLYLWVTEGKEAISINNTFLHL 178

Qy 176 GGLDIVHNHGDTYTRIPLFVOLFMPDNLYVPPFNTYFQSSSSGTTAR 235
 Db 179 GSDSKVTTGAEVTRIPVPLQFMPDERSVIAEPFNAHNSSIGKEKTYPLFFNEPFLNR 238

Qy 236 LIHDPCVIAFMVALVRVNTVAYARGAAHLLAFDENHEGAVLHRSIGEGFVYPPFNTYFGLCH 295
 Db 239 LLFBAVVGPAAVLCRNVAARAACHAFLFADITTAFAEASQG--KTPR 296

Qy 296 GARNDYNTSTSKPSKSPGCFERIASIMAADTAHAEVFNTGIVYETPTDIKEWPMFIGM 355
 Db 297 GER ---- DGGGGAAGSCFEQLASTMAGDALARLAEISIVSVAWFDEPPTDLSAWPFLFGQ 351

Qy 356 EGTLPRMLNALSYTARYAVGIVAMVPSNSALYLTEVEDSGMTEAKDGGGPPSPNRFYQF 415
 Db 352 DTAARANAVGAYLARAGLYCAMVFTNSNALHLETVDDAGPDPDKHSK-PSFVRFPLV 410

Qy 416 AGPHLAANPOTDRGHL----SSQSTSNSNTBSVYDYLALIGFGAPFLARLFLYI 468
 Db 411 PGTVAANPQVREGHVYPGPGFGRPAPLIVGGTQ-EFAGELAMLCGFSPLLLAKMFLY 469

Qy 469 ERCDAGAFTGGH-DALKYVTCEDSETPCSLCEKTRPYCAHTVHLRORMPREGOAT 527
 Db 470 ERDGAVTVGRREMDFVRYVADSNQTDVPLNCLTPTRHCVHTLMRLARHPFASAA 529

Qy 528 ROPIGVGFMTMSQYSDCDPGLNYAPYLILRKPGDQTEAAKATMDQTYRATERLRLFIDEQ 587
 Db 530 RGAIGVGFMTMSQYSDCDVLYAAFSALKR-ADGSEETARTIMQTYRATERVMAELET 588

Qy 588 ERILDRAAPCSSEGSSVIVDHPPTPERLDTLRARIEQTTFQMKVLYVETRDYKTRREGS 647
 Db 589 IQVVDQAVPTANGRLTITTRREALHTVNNVROYDREVBQMLNLYVEGRNFKERDGLG 648

Qy 648 ZATHSMALTFDYSGAFCPITFLVTRHILAVDODIALSQCCHVYGOVVEGRNFRNQFQ 707
 Db 649 ELNHANSILTDYFACFCPCPLIQLLGRSRNLAVYQDIALSQCCHVYAGQSVEGRNFRNQFQ 708

Qy 708 PVLERRRFVDFLPGGFISTRSITVTLSEG-PVSAAPNPTLQDAPAGRTFDDLLARVSVEV 766

SO	SEQUENCE	1196 AA	128423 MW;	A19CA844380DDTPS	CRC64;
Db	709 PVLRRVMDMFNSLSATLTLVALSEGAAICAPSLTAGTQTAPESSFEGDYARVTLGFP	766			
Qy	767 RDTRKXNTRVSGNCTNLSEAAARLVLGLASAORQERVDMLHGAJGFLLKQFGHLFP	826	Query Match Score: 49.9%;	Score: 3015.5;	DB: 1; Length: 1196;
Db	769 KELRYSRVLFLAGASANASEAAKARVASSQAAQPKPDRVILCPGLKLQXPHAAIPP	828	Best Local Similarity: 49.0%;	Pred. No. 1.7e-223;	
Qy	827 RGMPPNSKSDNPQWETLQLRNQMPADKLTHEBTTIAAVKRTEEYAAINFINLPPTCI	886	Matches: 595; Conservative: 213; Mismatches: 366; Indels: 41; Gaps: 16;		
Db	829 NGKPGSQNPQNPQWETLQLRNQPLPARLISREIETAPIKSLDGMFINLPNNV	888			
Qy	887 GELAQFYMANLILKYCDHSQYLINTLTSITGARRPRDESSVHLWIKDTSAADETOA	946			
Db	889 SELAMYMANQIIRYCDSTIFTNLTLIAGRRRPPEQVAAW--SAOGAGIEAGA	945			
Qy	947 KALLEKTENLPELWTATASTTHLYRAAMNQRPVIVGJISISCHGAGANNUPQANWQ	1006			
Db	946 RALVDAVAHGGTWSMASCNLRPMAAPRNIVLGSISKYGWAGNDRVQAGNWAS	1005			
Qy	1007 LNGKKNVCPLEPLFTDRTRFELTAPRGGBFICPYTCGPGSSGNRETTLSQVRGLIVSGAMVQ	1066			
Db	1006 LMGKRNACPLLIDPRTFVLAICRAGVCAASISLGGGAHEBSLCLQRLTISEGAAVY	1065			
Qy	1067 LAIYATVYRAVGARAQHNAFDWULSLDDEFLARDLLEBLHQIQTQLETPNTVEGAL--	1123			
Db	1066 SSVFVATVSKSLGRTRTOQIQEDWALLEDYEISSEMMELTARALENGNEMSTDALEVA	1125			
Qy	1124 -EVKVLDEKTTAGDGETPTNLAFNFD---SCPHSHDITNTNISNSNISGTSVPLKRP	1180			
Db	1126 HEEBALYISOLQNGAE-----VNFGDFGCE-----DDNATPFGGPAGPAFAGRKA	1173			
Qy	1181 PEDEDFPLSGIFKHGNITMEM	1203			
Db	1174 FHGDDPFG-EGPPDKKGBLTM	1195			
RESULT	6				
ID	DNB1_HSV2H	STANDARD;	ERT;	1196 AA.	
AC	P84452;				
DT	01-NOV-1997	(Rel. 35, Created)			
DT	01-NOV-1997	(Rel. 35, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Major-DNA-binding protein 8 (Infected cell protein 8) (ICP 8 protein)				
3N	DBP or UL29 OR ICP8.				
DS	Herpes simplex virus (type 2 / strain HG52).				
JC	Viruses; dsDNA viruses; no RNA stage; Herpesviridae;				
JC	Alphaherpesvirinae; Simplexvirus.				
JX	NCBI TaxID:10315;				
JP	SEQUENCE FROM N.A.				
JX	[1]				
RA	Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.				
JC	- FUNCTION: SINGLE STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA REPLICATION.				
JC	- SUBCELLULAR LOCATION: Nuclear (Probable).				
JC	- SIMILARITY: BELONGS TO THE HERPESVIRUSES DNA-BINDING PROTEIN FAMILY.				
DC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and/or commercial entities requires a license agreement (See http://www.isb-sib.ch/ or send an email to license@isb-sib.ch).				
JR	EMBL: Z86099; CAB06754; 1/-				
JR	Intra-Pro: IPR000635; Viral DNA-bind.				
JR	Protein: PF0074; viral DNA bp; 1.				
JW	DNA-binding; DNA replication; Zinc-finger; Nuclear protein.				
T	ZNFING 499 512 C2HC-type				
T	DOMAIN 1168 1196 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).				

2y	1013	VCP.LIFFDRTERPIIAACPRGGFICPVTGPSSGNRETTLSQVRGIIIVSGGAMYCLAIYAT	1072	Db	245	VGPAAVALRARNVDAVARAAAHLAFDENHEGAALPADITTAFAESQG--KPQEGRAR--	299
Db	1012	ACPLJEDTRKVEKIAACPRGRVCARSSLOGAHSLSQLRGIAEGGAAVASSVVA	1071	Qy	302	WNSTSKSPSPGERSLASMAADTALHAEVTFNTGYBTPPTDIKEWMFIGNEGTLPR	361
Db	1073	VVRAGSARAQHMAFDWLSLTDDFLARDLHDDQIOTLETPTWVEGL--EAFKI	1128	Db	300	--DAGNKGPGGEQRPLASMAQDAALETSIVSMAVEDPPDTTWELEGETPAAR	357
2y	1072	TVKSLGPRTQLOQEWLALIEDTLESEMEFTRALEGRGGWSIDALEYAEASAL	1131	Qy	362	INALGSYTARYAVGIGAMVPSNSAHLTEVDSCMTEARDGGCPSPTRFYQFAGPHIA	421
Db	1129	LDEKTTAGTAGDETTPTNLNAFNEDSCBSPSHDTSVNLNISGSNISGTTVPGJKRPPBDELDF	1188	Db	358	ACAGVAYLARAGLICAMVTSNSAHLTEVDGAPDPAKDHSK-PSFTRFLVPGTIVA	416
2y	1132	VSQLGAGE-----VNFGDFGDHDHASFGGLAAA-AGLAGVARKRAFPHFDDPG	1182	Qy	422	ANPOTDRDGHYLT-----SSQSTGSSNTBFSVDYLALICGGFAPLARLLFYLERCDAG	474
Db	1189	LSGIPTKHGNITMEM	1203	Qy	417	ANPQLDREGHVTPGYBGRPAPLVTGTQ-EFAGHLMICGFSALLKMLFYLERCDGG	475
Db	1183	-EGEPERK-K-DITLDW	1195	Qy	475	AFTGGHG-DALKYTGTFDSEIPOSLCEKTRPICAHTVHLRORMPRFGATRQPQGV	533
Db	476	VIVGROBMDVERYADSGQDVPCLNLCETETRHKFASAARGALGV	535	Db	476	VIVGROBMDVERYADSGQDVPCLNLCETETRHKFASAARGALGV	535
Qy	534	FGTMNSOYSDPDIGNYAPYLIRKPGDQEAKATMDTYRATLERLFDLQERLLDR	593	Qy	534	FGTMNSOYSDPDIGNYAPYLIRKPGDQEAKATMDTYRATLERLFDLQERLLDR	593
Db	536	FGTMNSAYSDDVLDIGNYAAFSALKR-ADDSENTRTRIMQETYRATERMYAELAQYDQ	594	Db	536	FGTMNSAYSDDVLDIGNYAAFSALKR-ADDSENTRTRIMQETYRATERMYAELAQYDQ	594
Qy	594	GAPCSSEGSISSVIVDHPTRRLDILTRARIEQTTOENKVLYVTRDVKIREGLSEATHSM	653	Qy	594	GAPCSSEGSISSVIVDHPTRRLDILTRARIEQTTOENKVLYVTRDVKIREGLSEATHSM	653
Db	595	AVPTAIGRLTIGFIREALHTYVNNIKQLDVREVEQMLRNLLIGRNFFERDGLAEANHAM	654	Db	595	AVPTAIGRLTIGFIREALHTYVNNIKQLDVREVEQMLRNLLIGRNFFERDGLAEANHAM	654
Qy	654	ALTFDPYSGAFCIPITNFYKTHIAVYDOLLSOCHCIFYGQOEGENFRNFOPVILRR	713	Qy	654	ALTFDPYSGAFCIPITNFYKTHIAVYDOLLSOCHCIFYGQOEGENFRNFOPVILRR	713
Db	655	SLSLDFTYTCOPCPULQOLLARRNVIQPLASLQCHGTYAGQVEGENFRNFOQPVILRR	714	Db	655	SLSLDFTYTCOPCPULQOLLARRNVIQPLASLQCHGTYAGQVEGENFRNFOQPVILRR	714
Qy	714	FVDLFNGGFISTRSITVITSEG-PVSADPNPTLQDAPAGRTFDGLARVSBEVIRDVK	772	Qy	714	FVDLFNGGFISTRSITVITSEG-PVSADPNPTLQDAPAGRTFDGLARVSBEVIRDVK	772
Db	715	VMDELNGFLISKAKTITLVAUSEGAACPSLTAACTOPAESSEFGDVAVRTLGFPKELRKV	774	Db	715	VMDELNGFLISKAKTITLVAUSEGAACPSLTAACTOPAESSEFGDVAVRTLGFPKELRKV	774
Qy	773	NRVYPSGNCNTNLSEAAARLVGLASAYQREKRVDMLGALGLLKOFHGLIPRGMPPN	832	Qy	773	NRVYPSGNCNTNLSEAAARLVGLASAYQREKRVDMLGALGLLKOFHGLIPRGMPPN	832
Db	775	SRLFAGASANASAAKARVSLQSAYQKPDKVTDLJLGPGLFLLKQFHAYTFPNGRPPG	834	Db	775	SRLFAGASANASAAKARVSLQSAYQKPDKVTDLJLGPGLFLLKQFHAYTFPNGRPPG	834
Qy	833	SKSENPQWWTLLRNQMPADKUTHEETTIAAVKRTTEYYAIAINFNLPPCIGELAQF	892	Qy	833	SKSENPQWWTLLRNQMPADKUTHEETTIAAVKRTTEYYAIAINFNLPPCIGELAQF	892
Db	835	SNQNPQWWTLLRNQNPALARLSREDIETAFIKRSLDYGAINFLNAPNVSELAMY	894	Db	835	SNQNPQWWTLLRNQNPALARLSREDIETAFIKRSLDYGAINFLNAPNVSELAMY	894
Qy	893	YMANLILKYDHSOYLINTLTSITGARRPRDDESVLHWIRKDVTSAADIEQAKALLEK	952	Qy	893	YMANLILKYDHSOYLINTLTSITGARRPRDDESVLHWIRKDVTSAADIEQAKALLEK	952
Db	895	YMANQILRYCDHSTYFINTLTVAGSRPPSVQAIAAWPO--GGAGLEGARALMDS	951	Db	895	YMANQILRYCDHSTYFINTLTVAGSRPPSVQAIAAWPO--GGAGLEGARALMDS	951
Qy	953	TENIPPELMTAFTSTHLYRAANKNORPMVYLGSISKTYGAAGNNRVEQAGNWSLNGGK	1012	Qy	953	TENIPPELMTAFTSTHLYRAANKNORPMVYLGSISKTYGAAGNNRVEQAGNWSLNGGK	1012
Db	952	LDAHPGAWTSMASCNLURPNAARPMVNLGSISKTYGAAGNDRVEQAGNWSLNGGK	1011	Db	952	LDAHPGAWTSMASCNLURPNAARPMVNLGSISKTYGAAGNDRVEQAGNWSLNGGK	1011
Qy	1013	VCLPFTEDRTRTFIACPRGGFCIPVTGPSSGNRETTSDQVRGIIIVSGGAMVQLAIYAT	1072	Qy	1013	VCLPFTEDRTRTFIACPRGGFCIPVTGPSSGNRETTSDQVRGIIIVSGGAMVQLAIYAT	1072
Db	1012	ACPLIFDRTRKVLACPRGGVCAASLGGCAHERSLSLDEFDLRDLFLLDQIQLTLPWVEGL	1071	Db	1012	ACPLIFDRTRKVLACPRGGVCAASLGGCAHERSLSLDEFDLRDLFLLDQIQLTLPWVEGL	1071
Qy	1073	WVRAVGARAHQMFDDWLSLTDDEFDLRDLFLLDQIQLTLPWVEGL---BAVKI	1128	Qy	1073	WVRAVGARAHQMFDDWLSLTDDEFDLRDLFLLDQIQLTLPWVEGL---BAVKI	1128
Db	1072	TVKSLGPQTQLOQIETMPLDEYVSEEMFTTRALEGHENSTDALEVAHEAEL	1131	Db	1072	TVKSLGPQTQLOQIETMPLDEYVSEEMFTTRALEGHENSTDALEVAHEAEL	1131
Qy	1129	LDKETTADGETPTLNIAFNDFCEPHTDTTSVNLNSGSNSGSTVPGLKRPEDELF	1188	Qy	1129	LDKETTADGETPTLNIAFNDFCEPHTDTTSVNLNSGSNSGSTVPGLKRPEDELF	1188
Db	1132	VSQLGRAGE----VNFGDFGDDHAAEGGLAAA-AAGANGVARGRAFHGDPPFG	1183	Db	1132	VSQLGRAGE----VNFGDFGDDHAAEGGLAAA-AAGANGVARGRAFHGDPPFG	1183
Qy	1189	LSGIPKGHNITMEM	1203	Qy	1189	LSGIPKGHNITMEM	1203
Db	1184	-EGPFKK-DLIDM	1196	Db	1184	-EGPFKK-DLIDM	1196
RESULT 8							
DNB1_HSVB2							
ID							
AC							
P12639							
DT							
01-OCT-1998	(Rel.	12	Created)				
01-AUG-1990	(Rel.	15	Last sequence update)				

DT	16-OCT-2001 (Rel. 40, Last annotation update)	
DE	Major DNA-binding protein.	
GN	Bovine herpesvirus type 2 (strain BMV) (Bovine mamillitis virus)	
OS	Bovine; dsDNA viruses, no RNA stage; Herpesviridae;	
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;	
OX	Alphaherpesvirinae; Simplexvirus	
NCBI TaxID	10296;	
RN	[1]	
RP	SEQUENCE FROM N.A. MEDLINE=8830631; PubMed=2841793;	
RA	Hammerschmidt W., Conraths F., Mankertz J., Pauli G., Ludwig H., Buhk H.-J., Buhk H.-J., Pauli G.,	
RT	"Conservation of a gene cluster including glycoprotein B in bovine herpesvirus type 2 (BHV-2) and herpes simplex virus type 1 (HSV-1)." Virology 165:388-405 (1988).	
RL	Virology 165:388-405 (1988).	
RN	[2]	
RP	SEQUENCE FROM N.A. MEDLINE=8830632; PubMed=247278;	
RA	Hammerschmidt W., Conraths F., Mankertz J., Buhk H.-J., Buhk H.-J., Ludwig H.,	
RT	"Common epitopes of glycoprotein B map within the major DNA-binding proteins of bovine herpesvirus type 2 (BHV-2) and herpes simplex virus type 1 (HSV-1)." Virology 165:406-418 (1988).	
RT	- - FUNCTION: SINGLE-STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA REPLICATION.	
CC	- - SIMILARITY: BELONGS TO THE HERPESVIRUSES DNA-BINDING PROTEIN FAMILY.	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to licensee@isb-sib.ch).	
DR	EMBL: M2163; A2A46051; - .	
DR	PIR: A29242; DNBBERG.	
DR	InterPro: IPR00035; Viral DNA bind.	
DR	Pfam: PF00747; viral DNA bp_1.	
KW	DNA-binding; DNA replication; Zinc-finger; Nuclear protein.	
FT	ZINC-FINGER	
SQ	SEQUENCE 1186 AA; 127285 MW; A5B6ECC1479BD2C CRC6;	
Query Match	46.6%; Score 2933; DB 1; Length 1186; Best Local Similarity 48.7%; Pred. No. 3.7-e-17; Matches 585; Conservative 188; Mismatches 373; Indels 56; Gaps 11;	
Qy	1 MENTOKTRV---VPTGLGVYACRVEDLDEIISPLAARTSDSDALLPLMRILVETV 57	
Db	1 MENKOKTATTVKVSPGPKVYARRPPCGETTEIISLARSADSPTVLLIGITVSSG 60	
Qy	58 FRSVSLAVSGARTGLAGAGITKLTSVHPSVTFHGKXKHLPSAAAPNLTAACNAR 117	
Db	61 EDYNNVAVVYGSRTVGGTGTGSKLMPKSHYAPSAYFGERHLAASSAANLSSLCDAR 120	
Qy	118 ERFGFSSRCQGPVDGAVETTGAICTRLGLEPENTILYVNTAIEKEAVMCMVPLHGG 177	
Db	121 VQFGFSSFKPKPCBAGETTGACLCHLGNPNBSLLMVIAEKFREAYTISNTLHGG 180	
Qy	178 LDIVHNHGDVTRIPFLPFVQFLMDYVNLVPDPFHNTHSRISGEFVPPFATNGLCHL 237	
Db	181 VGTVLAGEEVRRPQYPLQNFPMDCRAADPFNDRAIGEFTAYPUPFENAKLASU 240	
Db	DNB1_HSV1 STANDARD; ID 00344; AC 00344;	
2y	238 HDCVIAPMVALVRVNTAVARGMIALAFENHEAVLPDDITTYFQSSSNTTARGA 297	
Db	241 FGAAGVPAVALRARNYDVARA-AHLAFTDENHEGAALPADTTAFTDPGTS --- KAG 295	
2y	298 RRDNDNSTSKSPSGCFERRIASMAADTALHAEVTFNGTUYETPTDIKEWMFIGEG 357	
2b	296 HRNPREC-----GGSFEBQLASVMGDAALLSIMSMAVFEEPPTDIGTMELTCES 349	

Equine herpesvirus type 1 (isolate HV25A) (EHV-1); Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

PR5; 375 AA; DT 01-FEB-1994 (Rel. 28, Created); DT 01-FEB-1994 (Rel. 28, Last sequence update); DT 16-OCT-2001 (Rel. 40, Last annotation update)

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

JC Alphaherpesvirinae; Varicellovirus.
 DX NCBITaxonID=10327;
 RP SEQUENCE FROM N.A.
 UX MEDLINE=9406109; PubMed=8275122;
 IA Bell C.W.; Whalley J.M./
 IT "Herpesvirus ICPI8.5 and DNA-binding protein genes are conserved in
 equine herpesvirus 1."
 UL Virus Genes 7:219-228 (1993).
 IC -!- FUNCTION: SINGLE-STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA
 REPLICATION.
 IC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 IC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES DNA-BINDING PROTEIN
 FAMILY.
 IC -!- This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement. (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).
 R EMBL; D13930; BA03033; 1; -.
 R PIR; JQ0846; JQ0846.
 R IPD; IPD00635; viral DNA bp; 1;
 R PIR; JQ0846; JQ0846.
 R DNA-binding; DNA replication; Nuclear protein.
 R NON-TER 1 1
 SEQUENCE 375 AA; 40309 MW; ECF327925BBF993B CRC64;
 / Query Match 17.4%; Score 1094.5; DB 1; Length 375;
 Best Local Similarity 55.5%; Pred. No. 1.9e-76;
 Matches 212; Conservative 57; Mismatches 98; Indels 15; Gaps 5;
 / 828 GMPNSKGPNPWFYLQLRQNPADLTHETITIAAVKRFTEYYAAINFLPTCTG 887.
 / 2 GHPPGIDTPNPWFYLQLRQNPADLTHETITIAAVKRFTEYYAAINFLPTCTG 61
 / 888 ELAQFYMANLILKYCDHSQYINTLTSIITGARRPRDPSSWHLWIKDTSAADETQAK 947
 / 62 ELAQFYFANLVLYKCDHSQYFNGLTIAIVGSRRPDAAVLWDLRTINGAADEVEPAAQ 121
 / 948 ALLEKTEENLPETWTAFTSTHLVRALAMNQRMTVIGISIKYHGAGNRVFOQNWSSL 1007
 / 122 EVLQRUGSNPAWTGFTSTMNVRVYMDOREMVVIGLSIKYNGSAGNTRVFOQNWNL 181
 / 1008 NGKKNVCPPLPFDRTRPPIACPRGFICPVTGPSSGNRTEITLSPQRGLIVSGGANYOL 1067
 / 182 NGKKNVCPPLMFDRTRPPIACPRGFICPVTGPSSGNRTEITLSPQRGLIVSGGANYOL 241
 / 1068 AIYATVYRAVGARAQHMAFDWLSLTDDEFELDPEELHDQIQTLETPTWYGALEAKV 1127
 / 242 AVFAVTHALGARTQHLLAVDDWGLYDDELAASSDAUNATVVFQF-GEWSVAAQELVK 300
 / 1128 ILDEKTTAG--DGEGPTNLAENFDS---EPSHDITTSVNLNISGSNISGTSVPLKRPP 1181
 / 301 NNEAQITAGAAYAAGEG---AFDRAVCYGDTPQQSTSAAAPAGQKRSL 352
 / 1182 EDDELFDLSSGIPKRGNIIMEM 1203
 / 353 PDDLFDMGAPPEKSGLTPDM 374
 :SULT 10
 BI EBV STANDARD, PRT; 1128 AA.
 DB DNB1_EBV P03257; STANDARD, PRT; 1128 AA.
 QY 21-JUL-1986 (Rel. 01, Created)
 DB 21-OCT-1986 (Rel. 01, Last sequence update)
 QY 16-OCT-2001 (Rel. 40, Last annotation update)
 DB Major DNA-binding protein.
 BI BAL2.
 DB 316 RVAIHYRNASLAPIVSQTQFATNSVLV----SGVERKSTGGKKESELFNFSMTHGLGT 370
 QY 418 ----PMLAANPQTDRDGHVLSQSOSTSSNTFSVDULALICGFGAFLLLF 466
 DB 371 LQEGETWDPCRRPCRGSGWGGPDTG----TNGEHN--YAEVBLVVAASPSPNLLARYAY 422
 QY 467 YIERCDAGATGGHEDAL----KVTGTFSEIPPSLCLERCHTRPVCAHTVHRURQR 519
 DB 423 YLQFCQ----GQKSSLTPVPETGSYVAGAAASP-CSLCEGRAPAVCLNTLFFRLRDR 475

QY		520 MPRFGOATRQPQIGYFGTMNSQYSDCDPLGNYAPYLIL---RKPGDQTEAAKATMQDTYR 575
DB	476 FPPVIMSTORRDPVITSGASGSYNTEDFGNFINFDKEDDGOPPDRPYTWWLQNQLL 535	
QY	576 ATLERLFDLEQERLLDEGAPSSSEGSSVIVDPTFRLDTLRAIEQTITQFMKVLY 635	
DB	536 ERJSLRGIDAEGLKEKEPHGP-----RDPYKMEKDVADEAETVQFNMSMA 582	
QY	636 ETRDYKIREGLSEATHSMALTDPSGAFCPITNFKRTHIAVQDALLSCHCYVGQ 695	
DB	583 K-KNITYKDLVKSCYHNOYSCNPFAQCAPCIFTQLYRSLSLTIDISLPICMC-----636	
QY	696 QVEGRNFRNQFQPVLRREVDIENGGF----ISTSUTV----TILSEGVPVSAPNPLG- 745	
DB	637 -----TEND-----NPGLGSPPPEWIKGHQTLCTNPRSLAIDKGVLTAKEAKTVHGPICDl 689	
QY	746 -----QDAPAGRTRFDGDLPEAMSVVTRDIRKVNRYTFSGNCNTNSEAARLVLASA 798	
DB	690 PDLDAALQGRVYGRRLPYRMVKLMLPRNKIRNRVFTGENAQLNS-----FIKS 742	
QY	799 YRQEKRVDMLHGALGFLLKOFHGLLSPRGMPNPKSPNQWMWLLQRNOMPDK-LTH 857	
DB	743 TTRENYI--INGPYMKFLNTYHKLF-----PDTKLSSLYLWHNFSSRRSVPVSGASA 795	
QY	858 EBITTAA-VKRFTEYEAANFINLULPTCIGELAQTYMANILKCDHSQYLINTLTSII 916	
DB	796 EYSIDLALFDGDSRAHEEENYDLYVPEGNLYTYAKQRLNNAILKACGQTQFYISLIQGLV 855	
QY	917 -TGARRPRDPSVSLWIRKDVTSRADIEOAKALLEKTENLPETWTAFTSTHLRAAM 974	
DB	856 PRTOQSPARDPHVLG--TRAVESAA-----AYABATSSLTAT-TVVCAATDOLSQVC 905	
QY	975 NORPMVVLGISISKHGAAGNNRVFAGNTSGNGRNQVCPBLFTDFDRTRFFIACPRGG- 1033	
DB	790 KARPVVTLPTINKTGTGVNGNNQIFOAGN-LGYFMGRGV-----DRN--LLQMPGAGL 955	
QY	1034 -----FICPVTPPSSGNR-EFTLSDOVRGIIYVSGGAMV----QLAIAYAT 1073	
DB	956 RKQRGSSSMRKKKFVPAATPLGLTVKRTQTATTYIENTAGLEAIIISOKQBEDDVFDDV 1015	
QY	1074 VRAGVARAQHMAFDWLSLT-DDEPLARDIIEELHDQIQLTET-----PWTVEGALEA 1125	
DB	1016 CNLVDANGEEAA----SLTRDDAEFLLGFSVLDLSVLTATIASSGIEWTAEARDF 1070	
QY	1126 VKILLEKTTAGDGETPTNLAFNEDCEPSEHDTSVNLISGSNISG-----STVPEL 1177	
DB	1071 L---EGWNGGGPAACQDNFISVAEPVSTASQASAGLILGGGGQGGRRRKLATVLPJSI 1126	
QY	1178 K 1178	
DB	1127 E 1127	
RESULT 11		
D	DNB1_HSVSA STANDARD; PRT; 1128 AA.	
C	P24710; (Rel. 21, Created)	
C	01-FAR-1992 (Rel. 21, Last sequence update)	
T	16-Oct-2001 (Rel. 40, Last annotation update)	
E	Major DNA-binding protein.	
N	6 OR KFERF.	
C	Herpesvirus Saimiri (strain 11).	
C	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;	
C	Gammaherpesvirinae; Rhadinovirus.	
X	NCBI_TaxID=10183; [1]	
P	SEQUENCE FROM N.A.	
P	MEDLINE=9016321; PubMed=2154888;	
P	Albrecht J.-C.; Fleckenstein B.;	
P	"Structural organization of the conserved gene block of Herpesvirus	
P	saimiri coding for DNA polymerase. Clinico-virology. 1990; 10: 101-107.	

2Y 416 AGPHLAANPQTDRDGHVLSQSTGSSNTFSDVYDIALICGFAPLLARLFLYERCDAGA 475
 Jb 372 HGLAYAADTQRENGEPAFGAKPSGGTYTLVHLLASSFSPLLARNCYMFQFCQHQK 431

YY 476 FTGGGDAKLYVTGFDSEPPCSLCBKHTRPVCAHTTVERLROMPR-EGQATRPTGCFV 534
 Jb 432 STTNANVSPQQYVGAAASDPLCELQGTCFASCHTLFLRDLRPFPLGSQRDPYVT 491

Y 535 GTMNSQYSDDPLGNYAPYLRLRGDQTAATAKMDTYRATERLFLIDLEQERLLDRG 594
 b 492 G-VSQYNDLMLNFAPE--REKEDEAVQNAESEKTYWQLIQNVI---EKUSTMG 542

Y 595 APCSBEGLSSVIVDHPPTERRILRARIEQTITQEMKVLYVETRDYKIREGILSEATHSMA 654
 b 543 VTEGVG-SELITDQSFELKTFRDLDNVVDSEVWKFMNCVLVK-NINNFRETIKTVHVLH 600

Y 655 LTFPYSGAFCPITNLFLYERTHLAVYQDLAL-----SOHCFCFYQQVN 697
 b 601 YCCNVFWQAPCMMFLNFYKSVSLIQQDCIPAMTYEQDNPSIGNMPSENLVKHY-QT 658

Y 698 EGRNFR-----NOFQPYLARRFVDFLNGGFISTRSTRSLSBGPVSAPPNTLGQ 746

b 659 IWTNFKSSCLDRGYLTGSEHKVTHDMCDEFLN-----IDSALS-----GQ 699

Y 747 DAPGRRTFDGLARVSVEVIRDRVRVVFSCNTNLSEARARLVLGLASAYORQEKRV 806
 b 700 IVP--MKHQYRVLAKALLTVPKTIKRNIVSN--SSMTETIQSGFPIKSAT----RKDS 750

Y 807 DMHGALGEFLLKQDFGLARVSVEVIRDRVRVVFSCNTNLSEARARLVLGLASAYORQEKRV 864
 b 751 YIVTPYMKFLNSLHKYNE----PNAXISALYWHFTSQRKLPVLPGSRENMYELAN 805

Y 865 AYRFTEYYAAINFNLPLPTCIGELAQFYMANLILKYCDHSQYLINTLTSI-----ITG 918
 J 866 YVETSSKRHDMDANVLDIPTTLYAKVRLNTTILRIGQTYATTQCLPTQTTISA 865

Y 919 ARRPR--DPS--SVLHNIR--KD-----VTSADIEOTAKALILEKTENLPWTTAF 964
 J 866 TEYPHVLDQSMSVHDYLSSTKDKKHALTVCOTLREDIATVGK-----908

Y 965 TSTHLVARMORPQVFLGIGISKHYAAGNNRVFAGN---WSGNGGRNVCQPLFT-FD 1020
 J 909 -----QRPITVPLVYKTYKINGTQIFQCGNLGYFMGRGVDRLNLIPISTGGR 957

/ 1021 RTRRFLITACPRGFICP----VTPGSSGHR----TLSQVRSRIVSGAMVQLATYA 1071
 J 958 RONNSYMMRRHVEMPMVAHLVVKRSNLNLTFETETIRKVNQI(FEDD---NLUITD 1014

/ 1072 TV---YRAVGAORAHQAFDWLSTTDDEFLARDLHDIQIQLJETPTVGEALBAV 1126
 J 1015 NVVLELVKGLDSCENITEDDQFYGEYTIMSDEIWSRFQITSGAPSVENTKVL 1073

JRESULT 12
 JBI HSV6U STANDARD PRT; 1132 AA.
 JDNB1 HSV6U STANDARD PRT; 1132 AA.
 JP52339; 01-OCT-1996 (Rel. 34, Created)
 J01-OCT-1996 (Rel. 34, Last sequence update)
 J41. Human herpesvirus (type 6 / strain Uganda-1102) (HHV6).
 JViruses; ssDNA viruses, no RNA stage; Herpesviridae;
 JBetaherpesvirinae; Roseolovirus.
 JSEQUENCE FROM N.A.
 JMEDLINE=92666321; PubMed=7747482;
 JCompeals U.A., Nicholas J., Lawrence G., Jones M., Thomson B.J.,
 JMarston M.E., Bristathou S., Craxton M., Macaulay H.A.;
 JThe DNA sequence of human herpesvirus-6: structure, coding content,
 Jand genome evolution.";

RL Virology 209:29-51(1995).
 CC -!- FUNCTION: SINGLE-STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA
 CC REPLICATION.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- BELONGS TO THE HERPESVIRUSES: DNA-BINDING PROTEIN
 CC FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatioins Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X92436; CAA6167.1; -
 EMBL; X83413; CAA8375.1; -
 DR InterPro; IPR000635; Viral DNA bind.
 DR Pfam; PF00747; viral DNA bp; 1;
 KW DNA-binding; DNA replication; Zinc-finger; Nuclear protein.
 FT ZN FING 459 475 C4-TYPE
 SQ SEQUENCE 1132 AA; 127762 MW; 44663E3E559F0D8 CRC64;
 DR Best Local Similarity 7.8%; Score 492.5; DB 1; Length 1132;
 DR Matches 274; Conservative 185; Mismatches 544; Indels 253; Gaps 53;
 DR 2 ENTOKTITYPTPGLGVYACRVEDDLEBISLIAARSTDSDALIPLMRNUVKEPTSS 61
 Db 3 DENITVVSAPVSTAATWIVTPPKDKEELDVLSVLSMERNNSPPVISPLMNTVENDSTI 62
 DR 62 LAVVSGARTTGAGGITKLTKLTSHPVPSVTFVHGKHKV- PSSAALPNLTCACNARERF 120
 Db 63 V-----KTPUNNEGTTLTKTSEMPVCPETHGPEOLVGMAEBHGDLLRICEQTRKF 115
 DR 121 GSSECQGPVDGAVETGAEICTRGLBEPNTILYVVTALKEAIVFMCNVFLHYCGLDI 180
 Db 116 HQSFEVTVARYVIDIKA--LOSEAVKDDADSVTICHVAGNGPKEELF-----AGLI 165
 DR 181 -----VHINHGDVTRIPERPVOLFMDPVNRLVDPENTHRSIGEGTVYPTPFYNTG 232
 Db 166 PCYBEQIOVQVGBYSCVPLPSATLFE-----ETVLSSECTBTQERSEFLPA 216
 DR 233 LCHLIDCVIAMAVALRVNTAVARGAAUUAEDENHEGAVLPPDITTYFQSSSSGT 292
 Db 217 LSSTLFYVFTWGTTURFSNKTELDAGKROFTGEOTYKLAHKTYL-----266
 DR 293 TARGARNDVNTSKSPSGGERLARISIMADTALHAETFN-----TGIYETPTPDI 346
 Db 267 -- GISGOKISAVEKD-----FMLVDVSY-TV-TELSFHVAYLDSYD-PSQI 309
 DR 347 -- KEWMFIBGGTLPRNLAGSTTARAVGIGMVFSNSALYLTVE-----D 394
 Db 310 MNFDWPTIRNSETAHERMAQTLNKLHLSHLAVLIFANSILYCSKIAFIPNVKQAAN 369
 DR 395 SGTEAKDGGPSFSPRYOFGAHPLAANPOTDRDGH-VLSSQSTGSSNTPEPSYDYLALI 453
 Db 370 SMTQ-----ELLRLSLSLSCNAUSSLSDVYNDNRKIKCDTSGDKDKFSAHNLAYA 422
 DR 454 CGFGAPIALARLFLYLERCDAGAFTGHHDALKYMGTPOSEIPOSCLCEXHTRPVCAHTV 513
 Db 423 CTSPQPLSFVWNINRMSTYNAAGAHTEYNHLYNC--SANICBEFCDSKCCQCSCTGPM 480
 DR 514 HRLRQMPREQQATROPIGVFTGTMQSQSDCDPQNYAYPLURKPGDTEAKATMDDT 573
 DR 481 VRVGTLPALPKVNEKPSLWMSMFSRYAEVDLIGSF-----RKPVSELKEIGKDQNT 535
 DR 574 YRATLRLFLDQEBLRDLG-----PCSSGGLSSSTVDBHPTFR-----RILDT 618
 DR 536 LS-----LDRGKFVSIQPDYCKNSLIDBVGTGDTENVRSKKFDFVTSIHG 580
 DR 619 LRARIEQTTTOFMKVLYVETRDYKI-REGI-SEATHSMALTEPDPSGAFCPITNPLVKRTHL 677

581 LTCQCLEBCVS---RCIVEMRRTOTPREQENCILQSFNVDTPYATAFSPELTESYYKVL 637
 DR EMBL: AEP157706; AAB06339..1; -
 DR PIR: T44001; T44001.
 DR InterPro: IPR000335; Viral DNA bind.
 DR Pfam: PF00747; viral DNA DB; 1.
 DR DNA-binding; DNA replication; Zinc-finger.
 KW C4-type.
 FT ZN FING 459 475
 SQ SEQUENCE 1132 AA; 127817 MW; B3SAB4EAFAS103EFF CRC64;
 DR Query Match 7.8%; Score 490.5; DB 1; Length 1132;
 DR Best Local Similarity 22.0%; Pred. No. 3.2e-29;
 DR Matches 278; Conservative 180; Mismatches 535; Indels 273; Gaps 54;
 DR 2 ENTOKTIVTPGFLGYIYACRYEDLDLEELIFRLAATSTSDSIALPLMRNTIVTEKTFSS 61
 DR 3 DENEIVTSAVPSAAWIVVFPKBKELLDDVLSMERSNSVISSPILLMTLVENDFST 62
 DR 62 LAIVSGARTTGLAGAGITKLTTSHFYPNVFEGKHKVL_PSSAAPNLUTRACNARERP 120
 DR 63 V-----KPTINFGENTLKITISMPVCFFHGTEQLVMAEDGDLTRLCETOKF 115
 DR 121 GFCSRQCQPPVDPDAVEITGAETRIGLEPEPTILIVVTAFLKEAVFMCNVFLHYGLDI 180
 DR 116 HUGSFVEYETARKVTDIKA---LCSAVERKDADSVICHAVCGNGKFKE-----AGLLI 165
 DR 181 -----VHNGHDVDRIPLPVQLEMPDTRLYDPPNTTHRHIGEGIVVYPTPYNTG 232
 DR 166 PCVBEQIQVQGEYSCVKIPLYSATLPEFE-----ETISLSSCTEPQEROFFLPA 216
 DR 213 LCHLHDIVTAAAMAVALVRVNTAVARGAAHLAFDENHEGAALVPDIDTYPOSSSSGTT 292
 DR 217 LSRTLIFYVFWGTTLRFSTNTKELDAGKQFTQGBQTVKLADPHKTL-----266
 DR 233 TARGARDNDVNTSKSPSPCGFERRIASTVADTAHLAETVN-----TGIVBETPTDI 346
 DR 237 --GISQKISAVED-----FLMVDBSV-TEFSHYAELDSYD-PSQI 309
 DR 347 --KEWMFIGMEGTPLRNLGTSYTAVARAVGIGAMVESPNALYIYEV-----D 394
 DR 310 MNFNDWIAIRNSETHAERMAQTNUKHLSSHAVLIFANNSILYCSKULAPIVTKQAFN 369
 DR 395 SGMTEDGGCOPSFNFYQAGPHILANPQDTDRGH-VLSQSOGSSNTFSDYALI 453
 DR 370 SWTO-----ELLIILRSLSNCNASSLTDVYDNRKLIKCDSTSCKDKFSANHLAYA 422
 DR 454 CGFGAPILLFYLERCDAGAFTGGRGDALKYTTGTGDEI-----PSLCEBH 503
 DR 423 CATCPOLISYVWNLNRMSSV-----YRAGNAYTEYTHLWNCSANLQECDCXK 470
 DR 504 TRPVCAHTTIVRLRQMPRCQATROPIGYFGTMNSQYSDCDPQIGNAYYLIRKPGDGT 563
 DR 471 CCQSCIGTAMTRVGRFLPAFKVKEPLWMSMSRYYAEVDLGSFG---RKPVSL 525
 DR 564 RAAKATMQDTTRTLERFLDQEFLDRLORG-----PCSSCGLSVIVDIFPFR --- 613
 DR 526 KEGIDQONTS-----LDRGFYPSQIDFCRKSLSIDEVGTGEGTENTRS 570
 DR 614 -----RILDTRARARFOTTOEQMVKLVEPDYK REGISEATSMALTFDPYSGARFPI 667
 DR 571 KKDDEVSIHIGITOCIEECV-----RIVENRRTQPREQENCLQSENVDTTATARSFF 627
 DR 668 TNFLVTRTHLAVQDIALSOCHCVPYGOOGEGRMFNRQDQPVIRRREYDLF -----N 735
 DR 628 ITFAVYKVVITLQNL-----IVASGHVYDRPTGNSTSKWLVOQQYOSLYCPTFHSSVYK 683
 DR 720 GCFISRTSITLSECPVSAPNPTIGO---DA---PAGR---TFDGDIAVSYEVIRDIR 770
 DR 684 EGFLNTRTRVKG-----ASNDMEQIQLDCDLYKSGKXVTKTQIQLCBLNQCLDR 735
 DR 771 VQNRYVFSGNTNLSEAAARLYGLASA/QRQEKRVDMLHGLGFLKOFHGLFPRGMP 830
 DR 736 XANRPNNKSSATRHNYPKRNV-----REFRNPNPSGCSFLSKYHDRLF-----781
 DR 831 ENSK-SPNPPOWERFWTLLQRNQMPAKLTHEBTIT_IAAVKRTEYEAAANFINLUPPTCIGE 888

RESULT 13
 DNBI_HSV62 STANDARD; PRT; 1132 AA.
 AC P52578;
 DT 01-OCT-1995 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 JE Major DNA-binding protein (MDBP).
 IN U41 OR KA21.
 JS Human herpesvirus (type 6 / strain Z29) (HV6).
 DC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 DC Betaherpesvirinae; Roseolovirus.
 DX NCBI_TaxID:36351;
 RN [1].
 RP SEQUENCE FROM N.A.
 MEDLINE=95074921; PubMed=7983761;
 RA Stamey F.R.; Domingue G.; Black J.B.; Dambbaugh T.R.; Pellett P.E.;
 RT "Intragenomic linear amplification of human herpesvirus 6B oriLyt."
 J Virol. 69:539-596(1995).
 31 FUNCTION: SINGLE-STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA
 REPLICATION.
 32 SUBCELLULAR LOCATION: Nuclear (Probable).
 33 SIMILARITY: BELONGS TO THE HERPESVIRUSES DNA-BINDING PROTEIN
 FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions. There is no restriction on its modification and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announce/> or send an email to license@ebi.ac.uk).

782 PRVKISLEMLQRFLINNPVKTLIDGNPNEVKTKTFKAFASITNTYDEDDIIQPECLST 841
 889 LAQFYMANLILKYCQHDSQYL - INVLTS - LITG - - - - ARPRPDS - - SVLHWIRK 934
 842 FIDCYPHNKFSLALGFHDYTSLSHGUTSKVLTQNPVLFPVVLDKQPKESSIEBYLTVRK 901
 935 DVTSAADIEQTQAKALLEKTEBNLPPEDWTAAFTSTHLVRAAMNQRPMVYGLISKTYKHGAAG 994
 902 LVLDG - VPIVPIASISKEEN - - - FGTTFTS - - - - RSLVTFGLTLEKFTVSLA - 944
 995 NNRVFDAAG - NWSGLING - GRQNCPL - - - - FTIDRTRFETTACPRGGFICBVTGSSGN 1042
 945 NREYFOFGQLGWIGGGSVGDNLNPTSSALQDFRMFRM - ORTIAKTFSEVIV - - - KKVR 998
 1046 RETLTSQD - VRGILIVS - - - - - GGMAYQLAIATYVAVGARAQHMAFDMWLSLTD-E 1094
 999 RETIMFDTEVVKGVKVLISIVENLTDIDPLLLIAEVNR - - - - DREDKETMDMMLFVDGIRE 1055
 1097 FLARD - - - LEELHDQIQTLETPTWVEGALEYV - - - - - KILDEKTTSDG - - - 11138
 1056 ALAASTIMLKLHNHLDMMVTRDFESIA - NLQSVPFEAVSSNDAPVYDFSEELAEDDQASGVLK 1114
 1139 - EFTT 1142
 1115 CDETEET 1120

SPLIT 14

DBI_HS7VJ	STANDARD;	PRT;	1131 AA.
P55339;			
01-OCT-1996	(Rel. 34, Created)		
01-OCT-1995	(Rel. 34, Last sequence update)		
16-OCT-2001	(Rel. 40, Last annotation update)		
Major DNA-binding protein (MDBP).			
U4	Human herpesvirus (type 7 / strain JI) (HHV7)		
Viruses: dsDNA viruses, no RNA stage, Herpesviridae;			
Alphaherpesvirines; Simplexvirus.			
NCBI_TAXID=57278;			
SEQUENCE FROM N.A.			
STRAIN=J;			
Nicholas J.			
Submitted (JAN-1996)	to the EMBL/GenBank/DDJB databases.		
-!- FUNCTION: SINGLE-STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA			
-!- REPLICATION.			
-!- SUBCELLULAR LOCATION: Nuclear (Probable).	BELONGS TO THE HERPESVIRUSES. DNA-BINDING PROTEIN		
FAMILY.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: U43400; AAC54703; 1; -.
 PIR: "T41943"; T41943.
 InterPro: IPR00635; Viral DNA bind.
 Pfam: PF00747; viral DNA bp_1.
 DNA-binding; DNA replication; Zinc-finger; Nuclear protein.
 DNAbind; C2HC-type
 SEQUENCE 459 475
 ZNFINGE 1131 AA; 129008 MW; 52C97388D1B6D04P CRC64;

query Match 7.7%; Score 483; DB 1; Length 1131;
 best Local Similarity 20.8%; Pred. No. 1, 2e-28;
 matches 258; Conservative 204; Mismatches 560; Indels 220; Gaps 46;

2. ENTKQKTVTPGPGLGYVACRVEDIDLRRPSETAAARSMSMAYIIMNTTMYPPPMSC 51

Db 973 QDFRFMR-QKFVIATK---LCDI1-VKKVKRKEAIVDVEVINGKVNLNIESIISNSVNPE 1026
 Qy 1067 LAVIATVVAVGARAQHMAFDWLSLTDD-BFLARDLE-----ELHQJLIQETBP 1116
 Qy 1027 LILIAEVMK---DRDSKPQMMDMIFYVDRREPLAKSVKNIQLTDLNHDFSLSTLSV 1083
 Qy 1117 W-----TVEGALBEAVKILDEKTTAGDGETPT 1142
 Db 1084 FEEQEDSAIYDFSELLVEGNIEGGFLRKCESTEHEENEPS 1125

RESULT 15

DNB1_SCMVIC	ID _SCMVIC	STANDARD;	PRT;	1160 AA.
AC P13215;				
DT 01-JAN-1990 (Rel. 13, Created)				
DT 01-APR-1993 (Rel. 25, Last sequence update)				
DT 16-OCT-2001 (Rel. 40, Last annotation update)				
DE Major DNA-binding protein (MDBP).				
GN UL7 OR DBP.				
OS Simian cytomegalovirus (strain Colburn).				
OC Viruses, dsDNA viruses, no RNA stage; Herpesviridae;				
OX NCBII_Taxonomy: Cytomegalovirus.				
RN [1] SEQUENCE FROM N.A. PubMed=2172458;				
RX MEDLINE=91037979; PubMed=2172458;				
RA Anders D.G.; Gibson W.;				
RT Nucleotide sequence of a cytomegalovirus single-stranded DNA-binding protein gene: comparison with alpha- and gammaherpesvirus counterparts reveals conserved segments.";				
RT J. Gen. Virol. 71:2451-2456 (1990).				
RN [2] SEQUENCE FROM N.A. PubMed=98155776; PubMed=2831298;				
RX MEDLINE=98155776; PubMed=2831298;				
RA Anders D.G.; Gibson W.;				
RT Location, transcript analysis, and partial nucleotide sequence of the cytomegalovirus gene encoding an early DNA-binding protein with similarities to ICP8 of herpes simplex virus type 1.;"				
RT J. Virol. 62:1364-1372 (1988).				
JC - FUNCTION: SINGLE STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA REPLICATION.				
JC - SIMILARITY: LOCATION: Nuclear (Probable).				
JC - SIMILARITY: BELONGS TO THE HERPESVIRUSES DNA BINDING PROTEIN FAMILY.				

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

JC - EMBL: D00750; BAA00647.1; ALT SEQ;
 DR EMBL; M19868; AAA46166.1;
 DR InterPro: IPR000635; Viral DNA bind.
 DR Pfam: PF00747; viral DNA bp_1.
 CW DNA-binding; DNA replication; Zinc-finger; Nuclear protein; ZNFING.
 CW ZNFING_465 479 C2HC-TYPE.
 IQ SEQUENCE 1160 AA; 1129005 MN; 7037716816974B1A CRC64;

Query Match Score 7.6%; Score 479.5%; DB 1; Length 1160;
 Best Local Similarity 21.3%; Pred. No. 2.3e-28; Indels 255; Gaps 53;

Matches 279; Conservative 202; Mismatches 571; Insert 255; Gaps 53;

Y 1 MENTOKTIVPTGPGLGVYTACRVEDLDIBEISPLAARSTDSDIALPLMRNLTVEKIFTS 60
 b 1 MSNEBLSALAPVGRPAVYVFTKTHEMEVLATISLCDSSSPVTAAPLMGLTVDQDFCT 60

y 61 SLAVSGARTGLAGITKLTKLTHFYSVTFHGGRHLVP SSAAPNLTRACHAARER 119
 Db 1020 VLFDS---DIRCRVMALDSENLDVDEPELMANY---EILSTREEPERDVLFVFDGC 1072

Db 61 SV-----RTPWCYDG3VLTKVTS-FCPFALFYNTQGIVDFSEPHGDVQLCDETRQR 113
 Qy 120 FGFSRCQGPVFDGAVETTGAEICTRGLPEPENTILYVVTALFKEAVFMCMYFLHYCGGL 179
 Db 114 YATESM-----PEEGRAPTDLAALCTAACGDPQEYLVHVTGNGMKEFMYAGQ1LPCFBAA 171
 Qy 180 IVHINHGDYTRIPLPFPVOLF---MPDVNRLVPPDPFENTHRSIGEFTYPTPYNTGTLCHL 236
 Db 172 PTRLANDCDAVRVPYPPFLFGSLOADVD--SDELSLDKRS --SFVESRGLVYPAVSET 225
 Qy 237 IHDCVIAPKAVALVRVNTAVARGAAHLAFAFDENHEGAVLPDPDTTYFOSSSSGTTARG 296
 Db 226 LFVYVYTSMCQARFSETKVLTEAALQFVNDQSQSVCYKLAPEKKYGYTSOK-----277
 Qy 297 ARRNDVNSTSKPSKPSFGPERLASSIMAADTL-----HAVIENTGIYETPTDIKEWP 350
 Db 278 -----LSSIEKD-----HLMNSDIAVTCIELGFSFASVFLDSAYGASDSMTYSEWP 321
 Qy 351 MFIGMEGTIPRINALGSYTYARVAGIVAMVSPNSA-----YLTEVEDGMTEAKDGG 404
 Db 322 VVNTADHDLRALTEKLNLUHSITHSALLSCNSLYHNRLVLTSNKRNASGT----GA 377
 Qy 405 PGPSFNRYQFAGPHILLANPQIDRQH-VLQSQTGSNTNSFSDYLVATICGFGAPLLAR 463
 Db 378 SQEVLKLSIHFANGLTGLCETYDNDARKLIKSCGIVAKDERYAPHILSLICGTCQFLPSA 437
 Qy 464 LLVYLERCDAAGFTGGHGDALKYVTTGTFDS-----IPC-----S1.CEK-----HTRPVCA 509
 Db 438 FIVMLNRVS-----INTLGTGSSTLSHLIGLSSSLGACTGTCCT-----CY 482
 Qy 510 HTTVHRLRQRMQPRFGQATROPIGVFGTMNSOYSPCDPGLNYAPYLILRKPGDQTEA----565
 Db 483 NTAFVRVQTRLQMPRLPKRERPSVVMQSRFLNDYDVLGF-----RRYSAESKEASLDA 538
 Qy 566 -----AKATMDQTYRATLERLFDQERLDRGAPSSSEGSLISVIVD-HPT 611
 Db 539 KADEGSASTSNTASSVSDRTHR-LNRLDCKMRNLD-----SVTGBTDMTINGRSD 591
 Qy 612 PRILDITLARIEQTTOFMKVLTVERDVKIREGLESATHSMLTDFDPSGAFCPITNEL 671
 Db 592 FINLVSSLNKVKVDDEAMSFTS-----EVRMISNRNDEVLGATQAFNLDLNPAVSFFPILAYE 648
 Qy 672 VRTHLAVVQJALSOCHCVYQQ----QVEGRFRNQFQPVLRRFVDFLNGGFISTR 727
 Db 649 YYRVFAIIONVALITATSYVDNPLTTSLSVSRWVTHQFQI-----HGFSTTS 699
 Qy 728 TTVTLBEGPVAPNPFLQGDA PAGR-----TDFDLDARYSVEVIRDIRVK 772
 Db 700 ---RGKGLFIRNVKSSKRADHDLPDFKLYAEGTYSVIMEIKLSRSUVPSSLMFRVK 754
 Qy 773 NRVVFSGNCNTNLSEAAARLVGLASAYQOE--KRVDMLHAGLGFLLKOPHGLLPRGMP 830
 Db 755 NRP1-----SRAKSGT---TAHVFFERHEVKPNPKVKGCLGFLKYHDKLFP----799
 Qy 831 BNSKSPNPNQPOWFWLQLRQNOMPADLTHERTITAIAVRF-----TEYAAINFNINLPP 883
 Db 800 -DGFQSCUQFWQYKCANALP-----KVNKNIDGEFENNEFKVVISVADYNEDHDLIDVPP 853
 Qy 884 ICIGLAFQTMANLLRYCDSQL-----INTLTSITGARRPDRSSVHWRKDVTSAAD 941
 Db 854 DCMLNLYLENRFNKKLFCYGFKDQYIGTHLGLTRLYQHQAQF--YLGESMFNFAAAD 911
 Qy 942 BTQXAKALDEXTENPELMNTAFTSTLVRAMMNQRPMMVGLGISISKYHGAANNRVEQAA 1001
 Db 912 FAIRKLDL-KATGVTAPLASTTPTRESLTTFQRIQFVSIKVVNNKNEYQF 969
 Qy 1002 GN-----WSG-----LN-----GKRNVCPLPFTEDTRRFIACRGGFCICPVTGFSGNRBT 1048
 Db 970 QIGIYPSGNGVERSUNTSIGQD---YKFR-QRCLATKSDVLI---KRSRBDN 1019
 Qy 1049 TLSQYRGITIVSGGAMQ-----AIYATVRAVAGRAQHMAFDWLSLTD- 1095
 Db 1020 VLFDS---DIRCRVMALDSENLDVDEPELMANY---EILSTREEPERDVLFVFDGC 1072

Y 1096 EFLARDLEBLHDQQLIQTLETPTWVTEGALEAVKILDEKTTAGDGETPTNLAPNFDSCEPSH 1155
b 1073 QAVDSLMPKESKIQEMGUDDFSL--INLQQVLDSD&ECGGGEV-----R 1117
1156 DTTNVNLNTSGNSNTSGSTVPGLK-RPPD- -DELFDSLGIPIKHGN 1199
Y 1118 DL5ALFTAASGEAVGNSV -GLNARGGSHAFTD -DCGLLPKRGRL 1160

search completed: January 30, 2004, 13:13:04
ob time : 22 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

JM protein - protein search, using sw model

Run on: January 30, 2004, 13:10:47 ; Search time 27 Seconds
 (without alignments)
 4284.849 Million cell updates/sec

Title: US-09-769-699-2

Effect score: 6294

Sequence: 1 MENTQKTVPTGPGIYVIA.....DELFDSGIPKRGNTMEM 1203

Coring table: Blosum62

Gapop 10.0 , Gapext 0.5

Searched:

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database : PIR_76;*

1: Pir1;*

2: Pir2;*

3: Pir3;*

4: Pir4;*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query	Match	Length	DB ID	Description
1	6294	MENTQKTVPTGPGIYVIA.....DELFDSGIPKRGNTMEM	1203	100.0	1204	DNBEB29 DNA-binding protein
2	3620	5	5	75.5	1209	DNBEC4 DNA-binding protein
3	3583	5	5	56.9	1208	DNBEC4 DNA-binding protein
4	3054	4	4	48.5	1196	1 DNBEBK5 DNA-binding protein
5	3049	4	4	48.4	1196	1 DNBEBV1 Major DNA-binding protein
6	3047	4	4	48.4	1196	1 DNBEHF DNA-binding protein
7	3003	4	4	47.7	1197	1 A48350 DNA-binding protein
8	2933	4	4	46.6	1186	1 DNBEHG DNA-binding protein
9	1044	5	5	37.5	2 JQ0845 DNA-binding protein	
10	669	10	10	10.6	1128	1 Q0BE47 DNA-binding protein
11	633	10	10	1128	2 T42922 Major single-stranded DNA-binding protein	
12	618	10	10	9.8	1145	2 S55600 Single-stranded DNA-binding protein
13	604	9	9	9.6	1128	1 DNBEM1 DNA-binding protein
14	596	9	9	9.5	1127	2 T03105 Major single-stranded DNA-binding protein
15	490	5	5	7.8	2 T44245 EK506 polyketide synthase	
16	483	7	7	11.3	2 T41943 methyl-accepting chemotaxis protein	
17	479.5	7	7	11.6	2 A36256 internalin protein	
18	477	7	7	12.0	1 Q0BEW1 O6-methyl-guanine-DNA-binding protein	
19	412	6	6	11.9	1 A44051 DNA-binding protein	
20	237.5	3	3	4.8	2 S69894 Major DNA-binding protein	
21	137	2	2	15.04	2 T17425 ERK56 polyketide synthase	
22	120.5	120.5	120.5	1.9	2 AI3014 RINAGS-YTARVAGTIGAMPSNSALYEVTEVSGMFBAKDGCGPSPNRFQAFGPHL	
23	120.5	120.5	120.5	1.9	2 F98269 RINAGS-YTARVAGTIGAMPSNSALYEVTEVSGMFBAKDGCGPSPNRFQAFGHL	
24	119.5	119.5	119.5	1.9	2 AI1501 RINAGS-YTARVAGTIGAMPSNSALYEVTEVSGMFBAKDGCGPSPNRFQAFGPHL	
25	117.5	117.5	117.5	1.9	2 H83380 RINAGS-YTARVAGTIGAMPSNSALYEVTEVSGMFBAKDGCGPSPNRFQAFGHL	
26	116.5	116.5	116.5	1.9	2 H81702 RINAGS-YTARVAGTIGAMPSNSALYEVTEVSGMFBAKDGCGPSPNRFQAFGHL	
27	115.5	115.5	115.5	1.8	2 T17428 RINAGS-YTARVAGTIGAMPSNSALYEVTEVSGMFBAKDGCGPSPNRFQAFGHL	
28	115	115	115	1.8	2 B35963 RINAGS-YTARVAGTIGAMPSNSALYEVTEVSGMFBAKDGCGPSPNRFQAFGHL	
29	115	115	115	1.8	1 XYRTFA RINAGS-YTARVAGTIGAMPSNSALYEVTEVSGMFBAKDGCGPSPNRFQAFGHL	

30	114	114	114	1.8	1215	2 I52882 autoantigen - human hypothetical protein
31	114	114	114	1.8	1943	2 T23979 hypothetical protein
32	114	113.5	113.5	1.8	1886	2 S04921 nuclear pore protein
33	113.5	113.5	113.5	1.8	1682	1 C70588 probable mbtE protein
34	113	112	112	1.8	735	2 F87355 GGDGF family protein
35	112	111	111	1.8	1699	2 T14074 complement component
36	111	111	111	1.8	3255	2 G81702 adherence factor T
37	110.5	110.5	110.5	1.8	878	2 A87385 TonB-dependent receptor
38	110.5	110.5	110.5	1.8	2609	2 T40399 probable transport fatty-acid synthase
39	110.5	110.5	110.5	1.8	3069	2 H70656 protein tyrosine kinase
40	110.5	110.5	110.5	1.8	3069	2 S24103 probable integral protein
41	110	110	110	1.7	1039	2 AI0122 hypothetical protein
42	110	109.5	109.5	1.7	977	2 G72526 probable polyketide synthase
43	109.5	109.5	109.5	1.7	1501	2 T45623 hypothetical protein
44	109.5	109.5	109.5	1.7	1762	2 T03222 probable polyketide synthase

ALIGNMENTS

RESULT 1						
DNBEB29 DNA-binding protein - human herpesvirus 3						
C;Species: human herpesvirus 3; varicella-zoster virus.						
C;Date: 30-Sep-1988 #sequence_change 16-Jul-1999						
C;Accession: C27244						
R;Davidson,A.J.;Scott,J.E.;J.Gen.Virol.,67,1757-1816,1986						
A;Title: The complete DNA sequence of varicella-zoster virus.						
A;Reference number: A27345; MUID:8636657; PMID:3018124						
A;Molecule type: DNA						
A;Cross-references: EMBL:X04370; PID:95989; PID:CAA29121; PID:960018						
C;Genetics:						
A;Gene: 29						
C;Superfamily: herpesvirus DNA-binding protein						
C;Keywords: DNA binding						

Query Match	Score	DB 1;	Length
Best Local Similarity	100.0%	Score 6291;	Length 1204;
Matches 1203;	Conservative 0;	Freq. No. 0;	Indels 0;
Mismatches 0;	Gaps 0;		
1 MENTQKTVPTGPGIYVACVEDDLEEFSFLARSTDSDAILPLMNTVEKTFTS 60			
Db	1 MENTQKTVPTGPGIYVACVEDDLEEFSFLARSTDSDAILPLMNTVEKTFTS 60		
Qy			
1 SLAVSGARTGTLAGACTITKLTTSHFYPSTVFENGKRYVPPSSAPNLTRCAHARERF 120			
Db	61 SLAVSGARTGTLAGACTITKLTTSHFYPSTVFENGKRYVPPSSAPNLTRCAHARERF 120		
Qy			
61 SLAVSGARTGTLAGACTITKLTTSHFYPSTVFENGKRYVPPSSAPNLTRCAHARERF 120			
Db	61 SLAVSGARTGTLAGACTITKLTTSHFYPSTVFENGKRYVPPSSAPNLTRCAHARERF 120		
Qy			
121 GFSRCQQPPVDPDAVETGAETVAVEMCNVFLHYGGLDI 180			
Db	121 GFSRCQQPPVDPDAVETGAETVAVEMCNVFLHYGGLDI 180		
Qy			
181 VEINHGQVTRIFLPVQLFDVNLYPLDFTPNTGGLCHIHC 240			
Db	181 VEINHGQVTRIFLPVQLFDVNLYPLDFTPNTGGLCHIHC 240		
Qy			
241 VMAPVALYRVNTAVARGAAHLDENEGAVLPDITYTFQSSSGTTGARGARN 300			
Db	241 VMAPVALYRVNTAVARGAAHLDENEGAVLPDITYTFQSSSGTTGARGARN 300		
Qy			
301 DVNSTSISPSPGFGERLASTMAADTALEAVINTGYETTDIKPMPFMGEGTLP 360			
Db	301 DVNSTSISPSPGFGERLASTMAADTALEAVINTGYETTDIKPMPFMGEGTLP 360		
Qy			
361 RINALGSYTYARVAGTIGAMPSNSALYEVTEVSGMFBAKDGCGPSPNRFQAFGPHL 420			
Db	361 RINALGSYTYARVAGTIGAMPSNSALYEVTEVSGMFBAKDGCGPSPNRFQAFGPHL 420		
Qy			

09/769,699
 2-21-04
 Search Notes

421 AANPQTDRDGHVILSSQSTGSNTTERFSVDYLALICGFGAPLLFYLERCDAGAFTGCG 480

Db 421 AANPDTDRDHVHLQSQTSSNTTFSVDYLALCGFGLPLAFLFYERCDAGAFTGHH 480
 QY 481 GDAKYVGTGDEBILPCSCCEKTRPVCAHTTVERLORMPRGQATRQIVFGTMNSQ 540
 Db 481 GDAKYVGTGDFSEIPCSLCEKTRPVCAHTTVERLORMPRGQATRQIVFGTMNSQ 540
 QY 541 YSDCDPLGNYAPVLLRKPGDQEAAKATMDQTYRATERLRLIDGAPCSSE 600
 Db 541 YSDCDPLGNYAPVLLRKPGDQEAAKATMDQTYRATERLRLIDGAPCSSE 600
 QY 601 GLSSYYVDEPTFRILDILRAREQTTOQMVKLVETRDYKIREGLSEATHSMALTFDPY 660
 Db 601 GLSSYYVDEPTFRILDILRAREQTTOQMVKLVETRDYKIREGLSEATHSMALTFDPY 660
 QY 661 SGAFCPITNFLVKRTHLAVVODIALSOCHCVFYQQVEGRNFRQFOVPLRFEDLFNG 720
 Db 661 SGAFCPITNFLVKRTHLAVVODIALSOCHCVFYQQVEGRNFRQFOVPLRFEDLFNG 720
 QY 721 GFSTRSRTSVTLYSEGVPYASAPNPILQGQDAPAGTGFEDGLARYSVEVTRDLYKRNRYVFSGN 780
 Db 721 GFSTRSRTSVTLYSEGVPYASAPNPILQGQDAPAGTGFEDGLARYSVEVTRDLYKRNRYVFSGN 780
 QY 781 CTNLSEARARLVLGLASAYORQEKRVDMGLGFLIKQFLGLFRGMPNSKSPNQW 840
 Db 781 CTNLSEARARLVLGLASAYORQEKRVDMGLGFLIKQFLGLFRGMPNSKSPNQW 840
 QY 841 FWTLQLQRNQMPADKLTHEITTAIAVKRFTTEYYAAINFNLPPCTIGELAQFYMANLILK 900
 Db 841 FWTLQLQRNQMPADKLTHEITTAIAVKRFTTEYYAAINFNLPPCTIGELAQFYMANLILK 900
 QY 901 YCDHSQYI-LINTLISITGARRPRDPSSVLMWIRKDVTSAADIEQAKALLEKTNLPEIW 960
 Db 901 YCDHSQYI-LINTLISITGARRPRDPSSVLMWIRKDVTSAADIEQAKALLEKTNLPEIW 960
 QY 961 TTAFTSTHVRAMNORPMMVYLGISISKYGAAGNNRVEQAGNNSGLNGKNVYCPLFED 1020
 Db 961 TTAFTSTHVRAMNORPMMVYLGISISKYGAAGNNRVEQAGNNSGLNGKNVYCPLFED 1020
 QY 1021 RTRREFTACCPGGFICPVTGPSSGNRETTISDQYRGIIIVGGANVQLAIYATVRAVGAR 1080
 Db 1021 RTRREFTACCPGGFICPVTGPSSGNRETTISDQYRGIIIVGGANVQLAIYATVRAVGAR 1080
 QY 1081 AQHMAFDWLSLTDDEFLARLBLELDQIOTLTPWTEVGAALKIDEKTTAGDGET 1140
 Db 1081 AQHMAFDWLSLTDDEFLARLBLELDQIOTLTPWTEVGAALKIDEKTTAGDGET 1140
 QY 1141 PTNLANFDSCEP SHPTTSVNLINISGSNISGSTVPLGKRPDEDELFDLSGIPKHGNIT 1200
 Db 1141 PTNLANFDSCEP SHPTTSVNLINISGSNISGSTVPLGKRPDEDELFDLSGIPKHGNIT 1200
 QY 1201 MEM 1203
 Db 1201 MEM 1203
 QY 1201 MEM 1203

RESULTS 2

NBEC4

:Species: equine herpesvirus 1

:Note: Host Equus caballus (domestic horse)

:Accession: E36798

:Submitted to GenBank, March 1992

:Description: The DNA sequence of equine herpesvirus-1.

:Accession number: A36605

:Molecule type: DNA

:Residues: 1-1209 <TEL>

:Cross-references: GB:M86664; NID:9310791; PID:AB02466_1; PID:9330623

:Cross-references: GB:1831; MUID:9229566; PMID:1318606

:Contents: annotation: Possible protein-coding frames given

:Note: neither amino acid nor nucleotide sequence is given

:Genetics:

:Gene: 31

:Superfamily: herpesvirus DNA-binding protein

:Keywords: DNA binding

A;Title: The DNA sequence of equine herpesvirus-1.
 A;Relevance: Prod. No. 1e-266;
 A;Matches: 696; Conservative: 191; MisMatches: 239;
 A;Indels: 35; Gaps: 14;
 QY 1 MENTQKTVPTGPIGYTYACRVEDIDLEEISFLAARSTSDJIALPLMRNLTEVKFTFS 60
 Db 1 MESAPKTVSLPVSPLGYYYARQFASLQGTVSILTARSAVSVDSDLAVLPVIRGIVTEOFTT 60
 QY 61 SLAVSGARTGLAGAGITLKLTSHYPSVFPFHGGRHVPSAAAPNLTRACNAARERF 120
 Db 61 NVAVAGSXTGIGGGTGTKLUTPSHNPNAFVYGGSVIGASSNAPNLTRACEARRF 120
 Db 121 GFSDCQGPDPDVGAETVGAETTRLGEPENTILYVVTAAFLKEAVFMCNVFLHYGGLDI 180
 Db 121 GFSAFSSSPVNEAVETSGETCASLNSPETTALYVVTBPKEMVYVCTNPLHYGGST 180
 QY 181 VHNHGDVTRIPLPFVOLFMPDPYNRVLPDPFTNTHRSIGEGFVYPPFVNTGLCHLIHDC 240
 Db 181 VTIDGQDAMKIFPYQVOLMPYQVLMDFVNLASEPNAKHSIGDDFVYSPPFNDSLCLHKGY 240
 QY 241 VTAPEVALVRVRYNTAVARGAAHLAFDENHEGAVLPDPDITYFQSS--SSGTTTARGAR 298
 Db 241 VUGPAAVALVRVNLDGARGAAHLADENHEGSQLFDQDVFTFLPOSTOGNGKGSCRAQR 300
 Db 299 RDNVNSTSKBPSBGGTERPLASIMADTAIAAEVFTNTGIVETPDIDKQWMPFQMEGT 358
 Db 301 QGB-GSSGSKNASSGTERLAVSNMAADTALSVDLSIMAGIYDTELPVSWEWPVLSSGDDT 359
 QY 359 LRLNALGSYTRAVGUYGANFSPNSAUYLTEVEDSGMTBAKDGQGPSPSFNRVYQFAGP 418
 Db 360 -ESLEAIGAYAARLSCUVGAMFESANSVLYMTEVDDGGADGCGGS-NFSYHFRYLIAP 417
 QY 419 HLAANPDTDRGHVLT-----SSQSTGSSNTTEFSYDYLALICFGFAPLIAIRLFYLERCD 472
 Db 418 YVAGNPFDTKGRLVPLHTADQOAAPINGSNOEFSIDLALACGFCPQIARLFYLERCD 477
 Db 473 AGAFTG-GHGDALKVTVGTTDSEIPCSLCKHPTTRPVCAHTVHLRORMPFGQATRPI 531
 Db 478 AGTGFGERNETTALRILANTLESVDPGCLQTRPACAHITLHLRQLRPLFGAPVRAPI 537
 QY 532 GYFGTMNSOYSDCDPLGNTAYPLLRKPGDQTEAKATMDQTYRATERLRLFIDLEQERLL 591
 Db 538 GFGTMNSA-VSDCDVIGNTASYGALKRND-NEAPKSIMQDTYTRATMERLYVNLQOKLII 596
 QY 592 DR-----GAPOSSEGGLSSVIVDHPFTRRLLDTRARIETOTTOMKVLVETRDYKIREGL 646
 Db 597 DKEITLAQASPSAP-TSIVYHDQASFQIGLNSIKDTIGAAEAFMRVTLVEARDFKIREGL 654
 QY 647 SEATISMALTFDPYSGAFCPITNFVTKTHLAVYQDLSQHLCVCFVQVBEGRNFRNQF 706
 Db 655 ADANTMSISIDPSSSFCPVTSLPRLTSEGVPSAEPNPTLGQDAPAGRTFDGDLARVSEVI 714
 QY 707 QPVILRRFVDLNGFPISTRSITVTLSEGPVSAAPNPTLGQDAPAGRTFDGDLARVSEVI 766
 Db 715 CPVILRRFVDLNGFPIITAKTVVTVVSDGVLAPDLTRPASBPPTDQDGDAVSVBVL 774
 QY 767 RDIRYTNRYVFSGNCTNLSEAARLVLGASAYORQEKRVDMILHGAGELFLQFKGFLIFP 826
 Db 775 RDLYRKNTFLSNGANNSEAALARVAGMASAARRPDGSNTLNGAVFLYQYHGFLIFP 834
 QY 827 RGMPNNSKSPNPKQFWTLIQRNQMPADLTBEITIAAVKFTPEYAAINLNPPTCI 886
 Db 835 GHPEGIDTPNPKQFWTLIQRNQMPADLTBEITIAAVKFTPEYAAINLNPPTCI 894

887 GELAOFYMANLILKYCDHSQYLINTLTSITGARRRDPSSVLMIRKDVTSADIETO 946
 895 GELAQFVFAVLVLYKCDHSQFINGITAIVYSSRRDPAAVLAIDRTINGADVEPPA 954
 947 KALLEKTEBNLPEWTTAFTSTHLLRMMRNPVMVGLGISKYRAAGNTRVFGAGNW 1006
 955 QEVLRQIGNSNAAWTOCTFTSTINMVRMVDQRPMVYGLSISKYNGT 1014
 1007 LNGGKNCVCPLETFDRPRFLIACPRGFCICVGTGGSRRETTLSDQVRGIVTSSGGAYVQ 1066
 1015 LNGGKNCVCPMAFDRLRPRFLIACPRGFCICVGTGGSRRETTLSDQVRGIVTSSGGAYVQ 1066
 1067 LAYATVVRVAGRAQHMAFDDWLSUTTDEFLARPLDQLIQTLETPTWVYGALEAV 1126
 1075 TAVFAVLHALGARTQHLATDWLULVDEFLASLDALNATVQFGEWSYBAQFLV 1133
 1127 KILDEETTAG--DGESTPTNLNAFPDFSC--EPGHDITSVNLNUSGSNISGSVTPGKRP 1180
 1134 KNMFAQTTAGAAVASEG---AFFDGACVGDTPPQGSTSA---FNGGLAMAAAAPAGQERS 1185
 1181 PDDDFLDSLGIPKKHGNITMEM 1203
 1186 LPDDFLDFMGAPPBKPSGLFDM 1208

ESULT 3

NA-binding protein - equine herpesvirus 4 (strain NS80567)
 :Species: equine herpesvirus 4
 :Name: strain NS80567
 :Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
 :Accession: T42574
 :RefSeq: E.A.; Watson, M.S.; Perry, J.; Cullinane, A.A.; Davison, A.J.
 :Gen. Viro.: 79, 1197-1203, 1998
 :Title: The DNA sequence of equine herpesvirus-4.
 :Reference number: Z22173; PMID: 9603335
 :Accession: T42574
 :Status: Preliminary; translated from GB/EMBL/DDBJ
 :Molecule type: DNA
 :Residues: 1-1208 <TEL>
 :Cross-references: ENBL:AF0030027; NID:92605950; PIDN: AAC59547.1; PID: g2605975
 :Experimental source: strain NS80567
 :Genetics:
 :Gene: 31
 :Superfamily: herpesvirus DNA-binding protein
 :Keywords: DNA binding

Query Match 56.9% Score 3583; DB 2; Length 1208;
 Best Local Similarity 56.4%; Pred. No. 7.4e-264;
 Matches 689; Conservative 194; Mismatches 305; Indels 34; Gaps 13;

1 MENTOKTVTPGPGVYACRVEDDLIEEFLFLARSTDSDALLPLMRNLTYKETFTS 60
 1 MESAPKVKSLPVSPVSPGVYAAQNTVETEALTMARSTISDLAVLPVVRGLTYQTFT 60
 61 SLAVSGARTGLAGGITLKTTSHFYPSPVFVIGKHYLPSAANLTACNARERF 120
 61 NVAVAGSKTIGGGAGITKLTPSHFTPAVFTGGSVPGASSKAAPNTRACLARRF 120
 121 GFSRCQCPVPGVAVETGABECLTRGLPEPNTILYLVPAFKAVFMCNVFLHYGELDI 180
 121 GFSPFSSPPVVDNAVETSGEICASNLSPETTLLVVTETFKENVMTNFYGGNST 180
 181 VHINHGDVIRIPFLPVQLFMDVNRLVPDPFENTHRSIGEGFVPTPFNTGLCHLIHC 240
 181 VTHGQBAVKPIPYQLYMPDNVNLAAEFNSKRSIDDEFYSKPFNSDULRJGY 240

RESULT 4

DNEBKS

C;Species: human herpesvirus 1
 A;Note: host Homo sapiens (man)
 C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990
 C;Accession: A28601
 R;Gao, M.; Bouchey, J.; Curtin, K.; Knipe, D.M.
 Virology 163, 319-329, 1988
 A;Title: Genetic identification of a portion of the herpes simplex virus ICP8 protein rec

301 QGD-GSSIKNGSSSGIERLASSIMAADTALSVDVTEPEVSDUDPI-LSGVDDO 358
 359 LPLRNALGSYTARVAGYTGAVYSPSIALYTEVEDSGMTBEAKDGGPGPSFNRFYQFAGP 418
 359 REBLEAGAYASRLSLLGVANVTSANSVLNTEVDGGPADGKDAS-NPYHRYFLIAAP 417
 419 HLAANPZODRIGHVLS----SOSTGSNTTEPSVSYDYLALICGEGAPILLARYLFLYLERCD 472
 418 YVAGNFTQDKGRVLQHTADDPAAPINGSHOESFDYLAACGRGPOLLARILFLYLERCD 477
 473 AGAFTG GHGIALKTYTGTFSEBICSLCEKHTPVCACATTVHLRQRMFRGATROJ 531
 478 ATFGERNETALRYANTLSEVEPGCLTPATPACAHHTLHLRQLPRTGTPVRAPI 537
 532 GVFGTANNSOYSDCDPLGNAYAPYLILRKPDQTEAKATMQDTFATLFLIDEQEZRLL 591
 538 GIFTGMNSTSDCDVIGLNTASYGALKRPNP-NEAPKSTMQDTFATMRLVNDLQAKLI 596
 592 DRGA---POSSBGLSVITDHPFERRLUDTARIBOTTQEMKVLYVTDYKIREGL 646
 597 DEALAHAGTCOSAS-TGIVYKQDQSPNLISTKIDTEGAEQEMRILYVTDYKIREGL 654
 647 SEATHSMALTFDPYGAFCPTITNLVKRTHLAVYQDIALSOCHCFCYFQSQQVEGRNENQF 706
 655 ADANHTMSISDPPSSSFCDVPTSLSRRTIPTAQDLYLVSQZCQLFQGSVBRNENQF 714
 715 QPVUERRFYDLENGFPTSTRSITLSEGVPASAPNPILQGDAPAGRFDGDLARVSVEV 766
 715 QPVUERRFYDLENGFPTSTRSITLSEGVPASAPNPILQGDAPAGRFDGDLARVSVEV 766
 767 RDIRVENVYFSGNCNTNLSEAARLYCLASAYQERKVDMLHCGALGFLLKQPHGILPP 826
 775 RDLRKNRVLFSNGANSSEAARVAGMASYRPRPGSNLNGAVGFLYVQHKVLLPP 834
 822 RGMPNSKSPNPWOMTLLQNPARKLTHELITTAIAKPFTEAAINFNLPPTCI 886
 835 RGHPGIDTPNPONFWTLLQNPARKLSEKDIETTAIKEFSEHYSAINHEINTPNNI 94
 887 GELAQFYANLILKYCDHSQYLNLTISITGARRPDPSSYLHWKRKDVTSAIDETQA 946
 895 GELAQFYANLILKYCDHSQYLNLTISITGARRPDPSSYLHWKRKDVTSAIDETQA 954
 947 KALLEKTEBNLPEWTTAFTSTHLLRMMRNPVMVGLGISKYRAAGNTRVFGAGNW 1006
 955 QEVILQOLGSNPAAWTGPFASTNMVYRMDQRMWVIGLSISKYNGSAGANNRVRQANWNG 1014
 1007 LNGGKNVNCPFLFIDRTRFLZCPGEGFICVTPGSSGNRETTLSDQVRGUVLSSGAMVQ 1066
 1015 LNGGKNTCPMLAFDRTRFLZCPGEGFICVTPGSSGNRETTLSDQVRGUVLSSGAMVQ 1074
 1067 LAYATVVRVAGARACMAFDQDNLSTUDDEPFLARDLELHQDIIQTLTETWTTWCAEAV 1126
 1075 TAIVSVVLTALGARTORLAVDWLQVGDWLGIVDQVDTAISD1NALAAVVDQF-GWWSVREAQDM1 1133
 1127 KILDEBT----TAGGCBTPNLFNFDSECPHSHTTSNTNINISSNISSTVPLKRP 1181
 1134 RMDAQINMGMVYSTGDB----AFDGFACYGCDANQSSTEMGPA---SSAAPQGQKRPH 1185
 1182 EDEDELPLSGLFPIKGHNITME 1203
 1186 PDDILFDMGAPPEKSGLTFDM 1207

2Y	GGDIVVHNGDVIRIPPLPPVQLMFPDVNLVPDPFNTHRSLIGEGFVYPTPYNGLCH	235	RESULT 6
jb	GGSKVTGGAEVTRIPVVLQMLPDSRVIAPENANRHSIGEKFKTPPLPFNRPLNR	238	DNBEPHF
2Y	LIDHDVIAVALVRNTAVARGAAHLADEFHEGAVLPPDITYTYFQSSSSGTTAR	295	DNA-banding protein - human herpesvirus 1
jb	LLFEAVGPAVALCRNVDAAAGAAHLADEFHEGAALPADTPTFAEASGG--KPR	296	C;Species: human herpesvirus 1
2Y	GARRDNSTSKSPSGGFFRLASIMADTAHLAEVINTGYYETTDIKEWPMIGM	355	A;Note: host Homo sapiens (man)
jb	GGR----DGCGKAAGGFQRLLSVMAGDAALLESIVSMAVDEBPTDISAWPLPEQ	351	C;Date: 31-Mar-1990 #sequence revision 31-Mar-1990 #text_change 07-Jun-1996
356	EGTLBRNLALGSYTARVAGTVGANVFSNSPALSYLTEDSGMTAEKDGPGPSFENRYQF	415	R;Hammerschmidt, W.; Conraths, F.; Mankertz, J.; Pauli, G.; Ludwig, H.; Buhk, H.J.
352	DTAARANAVGAYLARAAGLVAVFTNSLALHTEVDAGADPKDHSK-PSFYRFFLV	410	A;Title: Conservation of a gene cluster including glycoprotein B in bovine herpesvirus t;
416	AGPDLAANPOTDRGHVL-----SSQSTGSSSTEFSYDYLALICFGCAPLLFLY	468	A;Reference number: A94381; MUID:88306231; PMID:2841793
411	PGRHAANPOTDRGHVWPFGFRGPTAPLUGGTC-EPAGEHLAMLCGSPALLAKMFLY	469	A;Accession: D9242
469	ERCDAGAFTGCHG-DALKXVTGTFDSE-PCSLCEKHPVCARTTVERLRQEMPRESQAT	527	C;Cross-references: GB:M21629
b	470 ERCDAGAFTGCHG-DALKXVTGTFDSE-PCSLCEKHPVCARTTVERLRQEMPRESQAT	527	C;Superfamily: herpesvirus DNA-binding protein
470	ERCDAGAFTGCHG-DALKXVTGTFDSE-PCSLCEKHPVCARTTVERLRQEMPRESQAT	527	C;Keywords: herpesvirus DNA binding
Y	528 RQPJGVFGTMNSQSDCDPLGNATAPYLILRKPSDQTEAKATMDQTYRATLERLFDLEQ	587	Query Match
b	530 RGAIGVFGTMNSMYSDCDVLGNAYAFAFSALKR-ADGSETPARTIMOETYRAATERVMALE	588	Score 48%; Best Local Similarity 49.8%; Pred. No. 4
Y	588 ERJJDAGCSCSEGLSSVYDHDPTERRIDLTUARIECTTQFMKVLYETDVKTRREGLS	647	Length 1196; Mismatches 199; Indels 49; Gaps 17;
b	589 LQYDQCAYPTANGRLETTINRREALHTVNNVYRQQVVREREVDLMLRNVEGNEKFDDGL	648	1 MENTOKRTVT--VTPGFLQYVV--ACRTEDDLIEEISLAAASTSDSALLDLMRLNLTVE
Y	648 EARTHSMALTFDPSGAFCPTINFLVKRTHLAVYQDIALSOQCHCFVPSQQVEGRNFNQFO	707	55; 1 METPKTATTIKVDPGFLQYVVARACPSBEGIE--LALLSASGDSDVAVLUVGIVTE
3	649 EANHMSLUDPAGCPBLQOLIGRSNLAYQDIALSOQCHCFVPSQQVEGRNFNQFO	708	58; 1 KTFESSLAVSGARTTGLAGAGITLKTTSHYPSVYRHGGCKHLPSSAAMNLTRACNA
Y	708 PYLRRRFYDLENGFISTRISITVLSSEG-PYSAPNPLGQDAPAGTFDGLARYSVBEVI	766	115 59 SGFEANAVVVGSTTGTGATVSLCQTPSHISSSTYVFRGGRHLDSITOANNLTRICER
3	709 PYLRRRFYDLENGFISTRISITVLSSEG-PYSAPNPLGQDAPAGTFDGLARYSVBEVI	766	116 ARERFGFCSRQGPBPVDFGAVETTGAEICTRIGLEPENTILYLVVTAFLKEAVFMCNVFLHY
Y	767 RDIKQKRVFVSNCNTLSEARALVGLASAYQORKVDMILHGAGLGELIKOFFGLLFP	826	117 117 118 119 ARRHEFGFSDYTRPFDGDLKRETCGEALCRLGIDPDRALLYAVVTEGEKEA,VSSINNTFLHL
3	769 KELRKVSRVLFGAGSANASEAKARYASLQSYQDKDVKVILLGPGFLLKQFFAAIFP	828	120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 199 200 201 202 203 204 205 206 207 208 209 209 210 211 212 213 214 215 216 217 218 219 219 220 221 222 223 224 225 226 227 228 229 229 230 231 232 233 234 235 236 237 238 239 239 240 241 242 243 244 245 246 247 248 249 249 250 251 252 253 254 255 256 257 258 259 259 260 261 262 263 264 265 266 267 268 268 269 269 270 271 272 273 274 275 276 277 278 278 279 279 280 281 282 283 284 285 286 287 287 288 289 289 290 291 292 293 294 295 295 296 296 297 297 298 298 299 299 300 301 302 303 304 305 306 307 308 309 309 310 311 312 313 314 315 315 316 317 317 318 319 319 320 321 322 323 324 325 326 327 327 328 329 329 330 331 332 333 334 335 335 336 337 337 338 339 339 340 341 342 343 344 344 345 346 346 347 347 348 349 349 350 351 351 352 353 353 354 355 355 356 356 357 358 358 359 359 360 361 362 363 364 365 365 366 367 367 368 368 369 369 370 371 372 373 374 375 375 376 377 377 378 378 379 379 380 381 382 383 384 385 385 386 387 387 388 389 389 390 391 392 393 394 394 395 396 396 397 397 398 398 399 399 400 401 402 403 403 404 405 405 406 406 407 407 408 408 409 409 410 411 412 412 413 413 414 414 415 415 416 416 417 417 418 418 419 419 420 420 421 421 422 422 423 423 424 424 425 425 426 426 427 427 428 428 429 429 430 430 431 431 432 432 433 433 434 434 435 435 436 436 437 437 438 438 439 439 440 440 441 441 442 442 443 443 444 444 445 445 446 446 447 447 448 448 449 449 450 450 451 451 452 452 453 453 454 454 455 455 456 456 457 457 458 458 459 459 460 460 461 461 462 462 463 463 464 464 465 465 466 466 467 467 468 468 469 469 470 470 471 471 472 472 473 473 474 474 475 475 476 476 477 477 478 478 479 479 480 480 481 481 482 482 483 483 484 484 485 485 486 486 487 487 488 488 489 489 490 490 491 491 492 492 493 493 494 494 495 495 496 496 497 497 498 498 499 499 500 500 501 501 502 502 503 503 504 504 505 505 506 506 507 507 508 508 509 509 510 510 511 511 512 512 513 513 514 514 515 515 516 516 517 517 518 518 519 519 520 520 521 521 522 522 523 523 524 524 525 525 526 526 527 527 528 528 529 529 530 530 531 531 532 532 533 533 534 534 535 535 536 536 537 537 538 538 539 539 540 540 541 541 542 542 543 543 544 544 545 545 546 546 547 547 548 548 549 549 550 550 551 551 552 552 553 553 554 554 555 555 556 556 557 557 558 558 559 559 560 560 561 561 562 562 563 563 564 564 565 565 566 566 567 567 568 568 569 569 570 570 571 571 572 572 573 573 574 574 575 575 576 576 577 577 578 578 579 579 580 580 581 581 582 582 583 583 584 584 585 585 586 586 587 587 588 588 589 589 590 590 591 591 592 592 593 593 594 594 595 595 596 596 597 597 598 598 599 599 600 600 601 601 602 602 603 603 604 604 605 605 606 606 607 607 608 608 609 609 610 610 611 611 612 612 613 613 614 614 615 615 616 616 617 617 618 618 619 619 620 620 621 621 622 622 623 623 624 624 625 625 626 626 627 627 628 628 629 629 630 630 631 631 632 632 633 633 634 634 635 635 636 636 637 637 638 638 639 639 640 640 641 641 642 642 643 643 644 644 645 645 646 646 647 647 648 648 649 649 650 650 651 651 652 652 653 653 654 654 655 655 656 656 657 657 658 658 659 659 660 660 661 661 662 662 663 663 664 664 665 665 666 666 667 667 668 668 669 669 670 670 671 671 672 672 673 673 674 674 675 675 676 676 677 677 678 678 679 679 680 680 681 681 682 682 683 683 684 684 685 685 686 686 687 687 688 688 689 689 690 690 691 691 692 692 693 693 694 694 695 695 696 696 697 697 698 698 699 699 700 700 701 701 702 702 703 703 704 704 705 705 706 706 707 707 708 708 709 709

Y	767	RDTVKNRYVPSGNCNTLSEAARARLVGLASAYORQEKRVDMLHAGFLKQFHGFLIP	826
b	769	KELRKSRLFLAGASANAEAAKARVASQSAVQDPRKDVTLLGPGLFLKQFHAIFF	828
b	827	RGMPNSKSPNPONFWTLQRONOMPDLTHTITIAVKRFTEEYAAININLPPTCI	886
b	829	NGKPGNSDNPONFWTLQRONQLPARLSREDETTIAFKESLDIGAINFINLPPNN	888
b	887	GELAQFYMANILKYCDHSQYLINTLTSITITGARRPRDPSSTLHWTRDVTSAADETQA	946
b	889	SELAMYNNQILRYCDESTYFINTLTIAIGSRRPPSVQAAA	945
/	947	KALLEKTENLPELMNTAFTAFTSLHYRAAMNQRPWVGLGSIISYHGAGNNRYFOAQNMSG	1006
b	946	RALYDAVDAHPGAWTSMPSACNLIRPVMARPKVYLGSISYHGAGNDRVEQAGNWAS	1005
b	1007	INGKGNVCPPLFTDRTRPFIACPRGGPTCPYGSSSNRENTLSDYRGTVSGCAMYQ	1066
b	1006	LMGCKNAACPILLIDTRKFVLACPRAFVCASSLGGAHESSLCLBURGTSEGAVAVA	1065
b	1067	LAYATVTRAVGARAQHNAFDWLSTLDEFIARLDEBLHDOTICLTTETPTVGEAL---	1123
b	1066	SSVFATVYKSLPRTQLOEDWLLADEYIUSEEMWELTAALEGNGENSTDALLEVA	1125
b	1124	-EAVKLDBKTTAGDGETPTNIAFNED--SCEPHSHTTSNVNINISSNSNISSSTYPCLKRP	1180
b	1126	HEEAALVSQLNAGE-----VNEDDFGCB-----DDNATPFCCGPAPPAFCKRA	1173
/	1181	PEDDELFLSGLSPIKHENITMEM	1203

SULT 7 8350

Human herpesvirus 2

Title: Nucleotide sequence of the major DNA-binding protein gene of herpes simplex virus. Reference number: A48350; MUID:9322841; PMID:895914
Accession: A48350
Modifications: DRY

Residues: 1-1197 <PDB>
Note: sequence extracted from NCBI backbone (NCBIN:129065, NCBIP:129070)

Superfamily: herpesviridae
Keywords: DNA binding; zinc finger

Query Match Score 3003; DB 1; Length 1197;
Best Local Similarity 47.1%; Pred. No. 2119.

Matches	Conservative	Mismatches	3'69:	Indels	40:	Gaps	16:
594;	4 TOKTIVPTGFLGKYY--ACRVEDDELEESIPLAARSTMSDIALPLAURNLTKEKTFRSS	212;					61

7 TTTTIVVPPGPMGYGRPAEGLL-USLISARSDADVAAVPLVLGVLTBESGPI 64
6 LAVVSGARTGAGAGITLKLTTISFYPSVVFVFGGGKHYLPSSAANPNTRACHAARRFG 121

65 VAAVYSGGTGGLGGTAVSCLMPHRYSPSTYVPGGRHLAPSTQAPNLTRLCERAHFG 124
 122 FSRCCOPPVGAVETTGARICTRIGRPERPNTIVVMMATPVEAVWMDPPTMCCGDTI 180

125 FSDYARPCDLKHETTGDAICERGLDPDRALLYVTEGFREACVTSNTFLRLGGDKV 184

182 HINRGGVIRIPLFPVQLEMPB^YNRLM^YDBBENTHEHSRISIGEGFVYPTPFNTGCLHLHDCV 241

RESULT 8

DNBEBG

Species: bovine herpesvirus 2	Strain: BMV
Date: 31-Mar-1990	#sequence_revision 31-Mar-1990 #text_change 07-Jun-1996
Accession: A29242	Title: Conservation of a gene cluster including glycoprotein B in bovine herpesvirus 2
Accession: A29242	Reference number: A94381; MOID: 88306231; PMID: 2841793
Molecule type: DNA	Cross-references: 1-1196 <HAMP>
Residues: 1-1196	Superfamily: herpesvirus DNA-binding protein
Keywords: DNA binding	
Query Match	46.6%; Score 2933; DB 1; Length 1166;
Best Local Similarity	48.7%; Pred. No. 2.2e-214;
Matches	585; Conservative 188; Mismatches 373; Indexes 56; Gaps 1
1	MENQKTVT---VPGTGPLGVYACRVEDDLEISPLAARSTDSIDLPLMRNLIVETKT 5
1	MENQKQTATVKVSPGPGLGVYARLPPGELTEALLSARSADSSTAVPLIAGITVESG 6
59	FITSSLAVVSGARTGTLAGAGGITLKLTTSHFYPSSVFHGGRHVLEPSSAAPNLTRACNAAR 1
61	FDNNVVAVVYVGSRRTGCGGTGVSVLKIMPISHYAPSAVVFHGGRHLAPSSAAPNLSSICDRAR 1
118	ERGFPSRSCQGPPVDAEVTTGAELCTRLGHPEPENTILYLVVTLFKEAVEMCYFLHYGG 1
121	VQFGFSSSPKPKPCEAEGETGELCAGLNPNESSLLVIAEGEKEAVIISTNLMEGG 1
178	LIDIVHINEGDVIRSIPLFLPVQLMFDPVNRLVPGAVETTGAEIOTRLGHPEPENTILYLVVTLFKEAVEMCYFLHYGG 1
181	VGTVTIAGEEVRIPIYIPIQMFMDYCAVADPFDNDRRAIGEYFAVPLPFVAKLASHL 2
238	HDCVIAPMVAALFVRVNTAVARGAAHLAEDNEHEGAVLPDITTYFQSSSSGTTTARGA 2
241	FGAVGPAVALRARNYDAVARAAAHAFDENNEGAALPADITTAFTDPTGS----TAG 2
298	RNDVNNTSKPSSEGGPFERRASIMAATTALHAEVIFNTGIVBTPDIKEWNVFIGMEG 3
296	HRNPREC----GGGFQRLASVYMGDAALALESMMSMAVEEPPTDGTWMLTCQES 34
358	TLPRLNAGLSTYARVAGVGTGAMYFSPNSALYLTTEVEDSGMTEARDGGCPSPSTRFYQFAG 41
350	TAARAASIGAYLGRAGLUGVAMYFSSNSALHTEVDDAGPADDXK -PTKPSFTRFLVPG 40
418	PHLAANQFTDRDGHVLSQS----TGSNTFSVDYLALICGFGAPILLARLFYLERC 47
409	TYVAANQDLDRGRVVAHGHEGPIVPGGNHFTCBELATLGFSPLLAKMYYLERC 46
472	DAGAFTG-GHGDALKYTTGTFDSEIPCSLCEKHTRPYCAHTTHTRLQRMPREGQATRQP 53
469	DGGVILGRPEMDFKYVSDAHTDVPCCLCS..LNREHSCAHTTILRLRARHPKTSITRGA 52
531	IGVFGTMNSQYSDCDPLGNYAPYLIRKPGDQTEAAKMDQYTRATLERLFLDQEQLRL 59
529	IGFVGWMNSAYEDCDVIGNYASFAISAIKRM-DVQETAAIMQETYRSAVERNMELINY 58
591	LDRGAPSSSEGLS.SVTDHPTFRRDLTLLRABORTQFMKVLTVERDYKIREGLSEAT 65
588	IDAAVITSPAKLESITITGREALQTV/SNVKQVDDGEVAQLMALVERGFRREALOEBAN 64
651	HSMALTFDPSGACPITNFVLYKRTHLLAVDQLALSQCHCVFYQQVEGRNPNQOPVIL 71
648	HAMSLLTLDPHASVPCPULLQMLGRNSLAVIQDQLALSQCHVYEGQAVEGRNFSQGPVIL 70
711	RRRFVDFLNGFISSTISITVLEGP-VSAINPNTLGODAPGRTFDODLARYSVEVTRDI 76
708	RRRVLDMFNNFGFLSARTLTYLDGACTISAPOLVSGQHAAAESGFEEDVAVNLGFPKEI 76
770	RYKQVVRVPSGNCTNLSSEAARARLVLGLASAYORQEKVDMJAGLGFLLQFHGLLPRRG 82

Qy	1182 EDEDELDLSGIPKHGNITMEM	1203	536 ERLSLGIDDEGLKEKEPHGP	582
Db	353 PDDILFDMGAPPERKSGL; FDM	374	QY	636 ETRDKIREGLSEATHSNAHTDPYSGAFCPTINFLYERTHLAVVODLALSQCHCVFYGQ
RESULT 10				
QQB47	DNA-binding protein - human herpesvirus 4 (strain B95-8)	C:Species: human herpesvirus 4, Epstein-Barr virus	Db	583 K-KNITYKDLVKSCTHVNOYSNCPFAQACPAFTQLYRSLLTILDQISLPCM-
C:Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 16-Jul-1999	R:Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.	Qy	696 QVEGRNFRNQFQPVLRREVDFLPNGK ---ISTRSTV----TUSEGPVSAPNPNTLG-	
R: Biol. Med. 1, 21-45; 1983	A:Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus	Db	745 YEND-NPELQGQSPPEWKGHQTLCNFRSLADGVLTAKTCVHGEPICDL	
A:Reference number: A91065; MUID: 85035713; PMID: 6032825	A:Accession: A43145	Qy	746 QDPAAGRFTFGDALARVSVEYTRDRVKRNRYVFSGNCNTNLSEAARARLVGLASA	
A:Molecule type: DNA	A:Residues: 1-1128 <BAN>	Db	749 PDLDAAALQGRVIGRRLPYRMSKYLCPRNKIKRNRYVFGENAALQNS-----FIKS	
A:Cross references: EMBL:VO1555; NID:959074; PIDN:CAA24808.1; PID:9134916	A:Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.	Qy	750 YORDEBKRTDMLRHEALGFPLKQFGLIPERGHPENSISSPNPQEWTFILQRQMPADK-LTH	
R:Baer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.; H	A:Reference number: A03305; MUID: 84270667; PMID: 6087149	Db	751 857 TTRRENYI--INGPMEFLNTRYTKLRL---PDTKLSSLYLWENFSRRREVVPVGASA	
Nature 310, 207-11, 1984	A:Contents: annotation; protein coding region	Qy	752 858 EETTTIAA-VKRFTEEEYAAAINFINLNPPTCIGELAQFYMANILKYCDSQQLINTLTSI	
A:Superfamily: herpesvirus DNA-binding protein	C:Keywords: DNA binding	Db	753 916 EETSDLAIFPDVGSRAHESNVYDWWEGNLTYAKQRLNNPKAGQTGYFISLIQGLV	
Qy	10.6% Score 669; DB 1; Length 1128;	Qy	754 917 -TGARRPRDPSSVLMWIRKQVTSADIEQTAOKALLEKTENPEWMTTASTTHLYRAAM	
Best Local Similarity 33.2%; Pred. No. 4.8%; Matches 292; Conservative 186; Mismatches 537; Indels 246; Gaps 45;	Db	755 917 -TQFQSPVARDPHVLG--TRAVESA-----AYABATSSLTAT-TVVCATAQCLSQVC		
Db	18 GPGGYYACRVEDLDLEBISFLAARSTDSDLALLPLMRNLTVKTFPTSSLAVISGARTG	Qy	756 905 KQAGQSSSMRKKPVFATPTPLQULTVVRTQATTBIEENRAGLEAIIISQREBICVFDVV	
Db	73 LAGAGITKLITSHSFPSVYFHGGKHLVPLSSAANLTRACNAERFGES-----RCQ	Qy	757 1034 -----FICCPVTGPSSGNR-ETTLSDOVRGIVSGGMV----OLAIYATV	
Db	75 DPNGL--LRAATSYHTRDIVFHNAHMVPPFFEGGLEALCGEREVEFGDAYSALPRES	Db	758 906 KARPVVLPVTKYGVNGNQIFAGN-LGYFNGRGV----DRN--ILQAPGAGL	
Db	127 GPPUDSAVETGAGLITRLGEPEVATVTLFKEAVFMNVFLYGLQDTHINHG	Qy	759 956 RQAGQSSSMRKKPVFATPTPLQULTVVRTQATTBIEENRAGLEAIIISQREBICVFDVV	
Db	132 SKPQDFPE-----GLDPASLYGAVAITAEFKELYSGNLVAAPSLSKQEVAVGQS	Qy	760 1074 VRAGVARAQNHAFFDDMSLT-DDEFLARDLIEBDHQIOTLET-----PTVTEGALEA	
Db	187 DVTRIPLEFPVOLMPDVNRLLYPDPFPNTHRSIGEFGVYTPPYNTGQLCHLIHDQVIAPMA	Db	761 1016 CNLVNDAGEA-----SLRDDAYLGLRFESVLAQDLSLTASSGIENTAAARDF	
Db	182 ASVRYPLVYDPEVLPVQD-----RQFINDSLRCMHEALYTGLA	Qy	762 1070 VILIDDEKTTAGDGETPTNLAFNEDSCEPHSHTTSVNLTNSGNSIG-----STVPOI	
Y	247 VALVVRNVTAVARGAAHLAFDENHEGAVALPPDITYTFOSSSSGTTARGARDNISTS	Db	763 1177 1-----L-----EGVWGGPGAAQDNFTISVABPVSTASQASAGLJGGQSGGRKRRLATVLSGL	
Y	223 QALRVRVGL-----VELLEKOSLQDQAKVAPKEFFAEST	Qy	764 1178 K 1178	
Y	307 KPSGGPFERRASIMAAAT-ALHAEVINTGIE---ETPTIK--BWPMPFGMBGTLP	Db	765 1126 D-----E-----EGVWGGPGAAQDNFTISVABPVSTASQASAGLJGGQSGGRKRRLATVLSGL	
Y	262 ISHPDSG-----AIMIVDSAACELAVASAPAMEASHTPASINYDSWPLADCPEEA	Qy	766 1127 E 1127	
Y	361 RLNALGSYTAARGVIGAMFSPNSALYLTEDVEGTMETAKDGGPGPSFNRFYQFAG---	Db	Db	
b	316 RVAALHRYNASLAPVHSTQFATFATSVL-----SGVKSSTGQKEEFLNFSTMHQLT	Qy	Db	
Y	418 -----PHIAANPCNDRDHVLSQGSTGNSNTFSVDYALICGFGAFLARLLF	417	Db	
b	371 LQEGETWDPCRRPCSGWGGPDVIG-----TNGPN--YAVELVYIASFSPLNARYAY	422	Db	
Y	467 YLERDAGAFTGGHDAL-----KYVTGTFDFSEIPCSLCFKHTRDVCASHVHLQR	575	Db	
b	423 YLQPCQ-----GKSSLTTPVPEPGTSYAGAAASPM-CSLCIGRARAFCVLTFLFLRDR	475	Db	
Y	520 MPRGQATRQPIGFGTMNSOYSDCDPLGNYAPYLIL---RKPGDOTEAKATHMDTYR	535	Db	
b	476 FPPMSTMSTQRDPYITSGASGNTNTDFQGNDPDKDDGQDPRPYTYWQNONQJL	535	Db	
Y	576 ATLERLFIDLEQERLLDRGAPCSSEGSSVIVDHPTFRILDRLARIEQTTQFMKVLY	635	Db	
Y	11 PTGLGYYVACRVEDLDLLEEISFLAARSTSDSIALPLMRNLTVEKTFETSSLAUVSSART	70	Db	

RESULT 11
T42922
major single-stranded DNA binding protein - ateline herpesvirus 3 (strain 73)
C:Species: ateline herpesvirus 3
A:Variety: strain 73
C:Accession: T42922
C:Accession: T42922
R:Albrecht, J.C.; Fleckenstein, B.
submitted to the ENBI Data Library, August 1998
A:Description: Primary structure of the herpesvirus atelieles genome.
A:Accession: T42922
A:Accession number: Z22274
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Sequence: EMBL:AT083424; PIDN: AAC95531.1
C:Superfamily: herpesvirus DNA-binding protein

Query Match 10.1% Score 633; DB 2; Length 1128;
Best Local Similarity 22.4%; Prod. No. 2.6e-39;
Matches 274; Conservative 194; Mismatches 514; Indels 240; Gaps 41;
C:Superfamily: herpesvirus DNA-binding protein

Query 11 PTGLGYYVACRVEDLDLLEEISFLAARSTSDSIALPLMRNLTVEKTFETSSLAUVSSART 70

Db 18 PVGAGTIVYVPEKGPFQEAISLLGNRNGADAMSILPLLSLTVEANFSENVKAVHKKID 77
 Qy 71 TGLAGAGITKLITSHYPSPVPHGGKHVLPSAAAPNLTRACHAARERFGFSRCQGPPV 130
 Db 78 M-----TTLSTVRSATREAVFENTNNEFFPPIPPGGLDSLCSARNLGYSTT-PRH 130
 Db 131 DGA---VETGAEICTRGLPEPENTILYLVVTALEKAEVMCNYFLHYGGLDIVHNG 186
 Db 131 DLKDLDIKDLYPAFTK---DSCFMSVTVTESKERLYFGNLVPLISQGQVQINGER 185
 2y 187 DTRIPFPVOLMPDFYNRLVTPDPENTHRSIGEGEVYPPFPNTGJCLHLHDVIAHPA 246
 Jb 186 EAVKIPLYDEDIFSKSHEHL---BPFYIHSVXLYLHDSELTSTA 226
 2y 247 VALRVNVNTAVARGAAHLAFTPDENHEGAVLPFDITYTYFQSSSSGTTTARGARRNDVNSTS 306
 Jb 227 QALRIRDVESTRAIRKQSTHDQYKLAKV-----VNSK- 259
 2y 307 KPSPSGGFERLASSIMAADTA----LHAEVIFNTG1-YEETP-----TDKEWMPF 352
 Jb 260 -----EPFAQVRCQDASAFWVDCFAELAISYGLSLEPQEQPCALLDTSWNPFI 311
 2y 353 IGMETGLPRNALGSTARAVGIVAMVEPNSAIYTIVEDSDSNTAEARDGGGPFNSRF 412
 Jb 312 DTTTEBEGRKIAQDNAMNSVHVYTHLTSVYLIRKINKQTOQNKSQDN--TINSY 368
 Y 413 YQFAGPHLAANPQDGDGHYLSQGSTGSNTERSDYDIALICFGAPILLALFYLERCD 472
 b 429 HQKSTTNNSVVAQVGTAVSECELCGKCPACHTLYKDRDPPVLISSQRDPY 468
 369 PMQHLSYAAEATQENQDPASGAVKENGSSYLFHALASSFSPHILARNCYMQFCQ 428
 473 AGATGGHGAAPDALKVVTGTFSEIIPSLCEKXHTREVCAHTTIVHLRQMRPFGAOATRQPIG 532
 Y 533 VFGTMNSQSDCDLIGNAYAPLIRKPGD---QTEANKATHMOTDTYATLRLFDLQE 588
 b 489 VVTGSGQNDLDMGNPATE--REKEDDMVNTCECKYTYQQLIONVVKL----- 538
 Y 589 RLLORGAPCSSEG----SSVIVDHEFTFRLLDTURARLEQTTFQMKVLYETRDYKIREGL 646
 b 539 -----ATIGITEGUGSOLSLITDQSEFEXTEREDDVDEVKFINLV-NNNINPRETI 592
 Y 647 SEATHSMALTFDPSGARCPITNLFLVRTHLAVVQDIALSOCHCVFYGQ----- 695
 b 593 KGVHVLHYCCNWFQACAMPFLNFKSVLAIQDCLPV--AMYEQDNPSMNMMPSE 650
 Y 696 -----OYEGRNFRN-----QFQPLIERRFVDLNGSFISTRSTIVLSEGPV 738
 b 651 WLKMHYQTWTNPKSSCDKGVLGASHKVVGDMPDFDINYD-----S 694
 Y 739 APNTLGQDAPAGRTEGDLARVSVEVDIRKRNRYFESGNCTNLSEAAARLWGLASA 798
 695 AIN---GQIVPV--KMQYRLLAKALITYPKTKKINTIVPSN--SNTTEAQGEF---- 742
 Y 799 YQREKRVD--MLHAGAFLKOFAGLFLPQRMPPNPKSNPQWFW-TLILQRNQMPA-DK 854
 b 743 -R-SIKRDKDSVYIWTGPMKFLNSLHWKIPF----DAKTSALYWHPSQRQPVLPG 794
 855 LTHEEITIA-AVKRFEPEYYAINFNLPLPTGIGLAFQYMANLILKCDHSQVINTL- 912
 2 795 ISKENNILEANIEAGSKMHDNVLDDIPTLAVRVLNNTLRLTCQTQYATTIQ 854
 Y 913 -----TSIITGARRPRDPSYSLHWKRDYTSAADETOAK--ALLEKTEENLPMLWTAF 964
 2 855 CLPLTQVSVATEYPH---VLA---QSITSVDDLISSINKQKAQIVQTLKEPATIG- 907
 Y 965 TSTHLYRAAMNCRPMVYLGISISKHGAGNNRFOAGN--WSSLNGKAVNCPLFT-FD 1020
 > 908 -----KRPITIVPLVNVKNTGINGNTQIQFCGNIGYPMGRGVDRNLTESSCFR 957
 , 1021 RTRRIIACPRGGFICPTVG----PSSGMRRET-----TLSDOVGTIVSGGM 1064
) 958 RONNSAYMRKHTIEMTPIVANLKTTSNLNLTFEVETIRKVNTQTFDDKDNLNFDNV 1017
 Qy 1065 VOLAIAYTVVRAVGARJOHMAFDWLSLTDDFLADLEHNDQIQTLETPTWTEGAE 1124
 Qy 1018 IEF-----VKLGIGESCNMPEDDLQXYLGYXYIMADEINRERFQLITDSCGPWSVTE--S 1068
 Db 1125 AVKILDEKTTAGDGETPTNLAF 1146
 Db 1069 VTKLIGSKK---QECQNLER 1085
 RESULT 12
 S55600 single-stranded DNA binding protein 06 - equine herpesvirus 2
 C: species: equine herpesvirus 2
 C: Date: 27-Oct-1995 #sequence_change 26-Aug-1999
 C: Accession: S55600
 R: Telford, E.A.R.; Watson, M.S.; Aird, H.C.; Perry, J.; Davison, A.J.
 J. Mol. Biol. 249, 520-538, 1995
 A: Title: The DNA sequence of equine herpesvirus 2.
 A: Reference number: 55554; PMID:9302501; PMID:7783207
 A: Accession: S55600
 A: Status: preliminary; nucleic acid sequence not shown;
 A: Molecule type: DNA
 A: Residues: 1-145 <TEL>
 A: Cross-references: GB:U00824; NID:965172; PID: AACI3793_1; PID: 9695178
 A: Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1995
 C: Superfamily: herpesvirus DNA-binding protein
 Query Match 9.8%; Score 618; DB 2; Length 1145;
 Best Local Similarity 23.1%; Pred. No. 3; 8e-38;
 Matches 295; Conservative 182; Mismatches 540; Gaps 47;
 Indels 260; Gaps 17
 Qy 4 TQKTVTVPTGPGVYACRVEDDUDLESISFLALARSTDDELLALPLMLNLTVKTKPTSLA 63
 Db 17 TQASV---GP CGPVLYLGGDTFVEESSLGNLHAGEEVFSPLSLLTVEADFRYVVK 72
 Qy 64 VUSGARTTGAGACTTLLTTSYFPTSYFVFGKHLVPSAATRCAACRAREFGES 123
 Db 73 -AVHKXLDPA----TVSYKASAYHREVIANAACPRKPIPGNGLCAASRQIFGYA 126
 Y 124 RCGCPVDPVDAVEETGTA----EICTRGLEP-EENTLYLTALPEAVMNCVPLHYG 176
 Db 127 EFE-----BZAGGAARPFELADGHLPBASHEAHGIVYTESPERLYRGQJYVVES 178
 Y 177 GLDTHYHNGDVRIRIPLEPVQLEMPDNRLLVEDPPNTHRSIGEGVYPTPYTNGLCLH 236
 Qy 179 QIQQVRVBGCDAAFPVYDGFELAKSPR-----ENTRY--PYHAGVSRV 221
 Qy 237 THDCAVIAPMAVLRVNTAVARGAAHLAFLDDNHEGAVLPPDITYTQOSSSGITARG 2 96
 Db 222 LFEAHYHTSLAQALRVDPGGLGALBROSP----HDQYKLPL--KVECREFPATGHRG 273
 Qy 297 ARANDVNSTSKPPGSGFERRASIMADTAHLAEVTFNTCTYBETP-TDKEVNFMFIGM 3 55
 Db 274 A-----GDSLSTTIVDWSATELLAVSYGLSFLVEQEGTALLSYDKNPIFEBC 3 19
 Qy 356 EGTLPRNLAQGXTARYAVGIGAMYVSPNSPASYTTEVEDSGMTEAKDGGGPSPTRYQF 4 15
 Db 320 EPEQRTEALTQNAOAVHVSQLSGNSTYLARVORKASN--RGCGCENVNTNSFNG 3 77
 Y 416 AGPHLAANPOTRDG----HVLSSOSTGSNTTEPSVDTIALICCGAPILLARLFLYER 4 70
 b 378 HGIACLALEPTQENGJLPSFPGPASALSGN--YSLHHLAYAASFSPOMLARICCYLOF 4 34
 Qy 471 CDAGAFQGHGALKVTFGDFDSEIPOSCEKRTPEVCAHTVIRLQOMPR-FQGATRQ 5 29
 Db 435 AQHQKSSNNSGNTVNPYVGTAAANTPMCELGRGSCASCNTFLYFLDRFPVYASVFRD 4 94
 Qy 530 PGVGFGMNSQSDCOPGLNAYAPYLJLREGDQTEAAKATMDOTYRATLRLTDLEOER 5 89
 Db 495 PYVTVGAGA-YNDLIDAGFNPY--RQDEDEIQSBEREKFTWQQTQVLR----ER 5 45
 Qy 590 LLDRGAPCSSEGLSSVTVDHPTFRILLTAR+BTOTTQFMKVLTETRDYKREGLESEA 6 9

RESULT 13
 JNBEM1
 DNA-binding protein - saimiriine herpesvirus 1 (strain 11)
 ;Species: saimiriine herpesvirus 1
 ;Note: host Saimiri sciureus (common squirrel monkey)
 ;Date: 31-Dec-1992 #text_change 16-Jul-199
 ;Accession: G36806
 ;Submitted to the EMBL Data Library, January 1992
 ;Reference number: A36806
 ;Accession: G36806
 ;Molecule type: DNA
 ;Cross-references: GB:X64346; PID:CAA45629.1; PID:g60327
 ;Albrecht, J.
 ;Virol. 66, 504-5058, 1992
 ;Title: Primary structure of the herpesvirus saimiri genome.
 ;Reference number: A37109; MUID:2333688; PMID:1321287
 ;Contents: annotation; possible protein-coding frames
 ;Genetics:
 ;Gene: 6
 ;Superfamily: herpesvirus DNA-binding protein
 ;Keywords: DNA binding

Query Match 9.6%; Score 604; DB 1; Length 1128;
 Best Local Similarity 22.9%; Pred. No. 4.3e-37;
 Matches 275; Conservative 186; Mismatches 509; Indels 226; GePs 40;
 Qy 11 PTGPUGYVIACRVEDLDLEETISFAARSTDSDIALPMLNIVTEKFTTSIAVSGART 70
 Db 18 PVEPCGVIYYPKEGFPEKASLLIGNKNGASAMSPLSLTIVESNSFNVKAHKID 77
 Qy 71 TGLAGAGITKLITSHFTYPSVFHGGKHLVLPSSAAPULTRACNAAEERFGSRCQPPV 130
 Db 78 M----TMILRVSAVREAVTFNTDLEPFLVGRDLIICSDARSFLGTNF--VPR 129
 Qy 131 DGAVENTGGAICTRGLBENTLYLVTALECAEVMCNVELHYGLDWHINQGDVI 189
 Db 130 TDLRDTVIDKLIVAFYSSDSCPNMVWTEGFKERLFGNLNLYPIIAQLK-VQINGREAV 188
 Qy 190 RPLFLPVQVLFMDPYNRVLPDPENTHHSIGEFPVYPPPFNTGCLHLHDCTIAPVAL 249
 Db 189 KPLPDEDIFSKSHEHL---PRFVLPSPVSKYLHDSDVFTSIAQL 229
 Qy 250 RVRVNTAVRGRAHIAFENHEGAVLPDITITYFQESSSGTTARGARDVDSNLSKPS 309
 Db 230 RIRVEYSTRASKQSITODQYKIAKI-----VNSK --- 259
 Qy 310 PSSGFERRLASIMAADETA----LHAEVLENTGII-YEETP----TDKEWMFIGM 355
 Db 260 -----DESQIUSKQCQDASAFYIDCIAELMSYGLSFLERPDQDP-AVLDUTSWLFLETA 314
 Qy 356 EGTLPRNLALGSYSTARAVGIGAMVSPNSAALTYTEDSCMTAEKDGGSQPSFRYQF 415
 Db 315 FTEEDRKAQDNNAMSVATHLSTNSTLTKINKQQQNSEQN--VNTYEMQ 371
 Qy 416 AGPHLAANPOUDRGHVLSQSTGSANTSESYDVYALICFGAPAPLRLFYLERCDAA 475
 Db 372 HGLAYAADATQENGESFASGPKFEGGTLYHALASSSPHILARNYCMFCOHOK 431
 Qy 475 FPGCHGDAKYTGVTGTPDSEIICSLCEKHTKRCVACTTVAIRLQRQMPR-EQATROPIGVF 534
 Db 442 SITNNATSVPVQVGTAAASDCELCOGTCPSCHTLFLYRKLDFEPVPGSQRDPPVYT 491
 Qy 535 GMNSQYSDCOPDGLGNYAPYLRLRKPDQDTEAKATMDYTRAUERLFLDEQRLDRG 594
 Db 492 G-VSGQNDLMLGNPATE--REKEDEAQNAESEKTYWQLQNYU--BKLSITMG 512
 Qy 595 APQSSEGSISSIVDHTHTFRLDTURARTEQTTOFMKYLVETDYKREGSEPAHMA 654
 Db 543 VTEGTG-C-SELIDT-QSPLKTFRDIDNVDSEVTKFMKLNVK-NINNFRTEITKVHVLH 600
 Qy 655 IATFDPSGAFPITPLVKEFHAYQDIAL-----SOCHCYFYGQV 697
 Db 601 YCCNVPMQAPCAMPLANLYFKSVLAIIQDCIPLANTYEDQNPNSIGMMSEWKVHY-QT 658
 Qy 698 EERNFR-----NOQOPVLRRFYDLEFNGFISTRSTIVTSEGPVBSAPNPTIGQ 746
 Db 659 TWNFNSCLORGVLTGSEKRVITDMFCDFLN-----IDSALS-----GQ 699
 Qy 747 DAPAGFTEDGHLARYSVEVIRDIVKRNIVFSGNCNTNLSEARARVGLASAYQRQEKV 806
 Db 700 IVP--MRMOTVLAKALLTVPTKIKTRKRVNFSN-SSMTETOSGFIKSAT----KKDS 750
 Qy 807 DMHGALGFLIKQFQGLLFPGMPPNSKSPNPQFWTLQRNOMPA-DLTHBEITIA- 864
 Db 751 YIVTCGYMKELNSHKVMP----PNAKISALYWHTSQKQLPVLPGISRENVELAN 905
 Qy 865 AVRTEEYAAINFNLPTCIGLAFQMFANILKYCBHSPQTLINTLTSI-----ITG 918
 Db 806 YVETSSKMHDDMNDIIPITLTYAKTRLNTLIRLTCQATTQCLPLQLTQTS 865
 Qy 919 ARRP-----DES--SYLHNTR-KD-----VTSAADETQAKALLEKTENLPDWTTAF 964
 Db 866 TEPAVLLDOSINSVDRLLSISKDKHALTVQTLKEDIATVGR----- 908

Qy	965	TSTHLVRAAMNORPMVYVIGISISKYHGAAGNNRVRQGN---WGGUNGKNCVCPLET-FD 1020	Db	496	TNVENELDFLGINFASFDKDEONQTE - ETPRPTWQLMQTLPKLEAGLVD - SPV 550
Db	909	-----ORPIVTPLYNKYTCTINGNTQIQFGQNLGYEMGRGVDRNLIPDSTGFR 957	Qy	598	SSEQL-SEVIVDEPFTFRILLDTLARIEOTTTOFMKVLTREYKTRGATISMAIT 656
Qy	1021	RTRFETACPRGEGFICP----VTPGPGSGNETB---TLLSQVRGLIVSGGAMVOLAIYA 1071	Db	551	ADEGAGGGSANNELEKFVTFPSDLSVDAEARFKINTMIK-NVNRKESIGVSVIQQN 609
Db	958	RQNNSSYMMRREHFVMTMVAHLVKNSNNNLTVEVITRQNQNPEDD----NLNIFD 1014	Qy	657	FDSYSGAFCPCITNFLVERTHLAVYDODILSCHCVPYGQQEGRNERNQGP - VLRRF 714
2y	1072	TV---VRAVGARAQHNAFDDWLSLTDTDLEBLHDQIOTLETPWVGEALEAV 1126.	Db	610	CNTYIWQADCSLMNLYVRSITIQLDAP----ISTYSESENPAOGQKPNVLUKYH 663
Db	1015	NVYLELVKGDBSCENTEDDLQFYLGEYIYIMSDIEWSRFQITDSGAPNSVENTVKV 1073	Qy	715	VDLFNGGFISTISITVLSGPVSAPNPLQODAPAGRPTGKD----LA 759
RESULT 14					
T03105 major single-stranded DNA binding protein - alcelaphine herpesvirus 1					
C;Species: alcelaphine herpesvirus 1					
C;Date: 24-Mar-1999 #sequence_change 24-Nov-1999					
C;Accession: T03105					
R;Ensler, A.; Fleckenstein, B.					
J;Viroi, J.; 6517-5525, 1997					
A;Title: Primary structure of the alcelaphine herpesvirus 1 genome.					
A;Reference number: Z14840; PMID: 97404659; PMID: 9261371					
D;Accession: T03105					
A;Status: preliminary; translated from GB/EMBL/DDJB					
A;Molecule type: DNA					
I;Residues: 1-1127 <ENS>					
;Cross-references: EMBL:AF005370; NID:92337967; PID: AAC88057.1; PID: S23337973					
;Superfamily: herpesvirus DNA-binding protein					
;Keywords: DNA binding					
Query Match 9.5%; Score 596; DB 2; Length 1127;					
Best Local Similarity 22.8%; Pred. No. 1-7-36-36					
Matches 281; Conservative 209; Mismatches 542; Indels 202; Gaps 43;					
Y 11 PTGPGUYTACRVEDLDLBEISFAAARSSTDSDLALLPLMRNLTVEKTFSSLAWSGA- 68					
b 20 PIGPGCIVIYPETHFNPKISLIGKDYEAAGAISPLPLRGTVEEAVPNTKAVERKID 79					
Y 69 RTTIGLAGAGITKLTTSHYPSVYVHGCKHVLESSAAAPNTTRACHNARERGFSRQGP 128					
b 80 MTT----VSKVJST--SYNRAIIFPVYKEFESIIFSQRGLGSLKREACDLFGY--VPFT 129					
Y 129 PYDGAETTGAETCTRIGEPENTYLVLVTAALKAEVAFMCNVFLHYGGLDIVHNGDV 188					
b 130 PL-SERGSTDVSIDC1BPVHQE-KDAIMAVVITEFKERHLGLKLIXLQSOMSVINKTEV 187					
Y 189 IRIPLFPVQMPDNVRLVPDPPTTHRSIGEGVYPTPFTNTGCLHLHDVIAANAVA 248					
b 188 YRFLPYDBLF----TKSSLR----LYLPAVSBELYTLTSLAQ5 227					
Y 249 LRVNNNTAVARGAAHLAFDENHEGAFLVLPDITYTYFQSSSGTTARGARNDVNNTSKP 308					
b 228 LRVNNNAASIVEAQCQYHDKYKMAKLYSPKEPLATVGAQDTTL-----272					
Y 309 SPSGGFERPLASIMAAADTALHAEVIFTNGIYEFPT-DIKEKPMFIGMEGTIPRNALGS 367					
b 273 -----MVIDAVAAEGLSISLSFFFBAQKTKVQDYSWDFASCETSDRLEALK 324					
Y 368 YTARVAGVIGAMYFSPNSALYLTVEVEDSGMTAKDGPGPS----FURYQAGPHL 420					
b 325 WNAQAIITHAQLFSTIYYNVARQ----APIPNISKVPEVNNSYVQHGLAN 376					
Y 421 AAQFQTDRDGHVLSQSQTGSS---NTEPSVSDYIALICCFGAPLARFLYLERCDAGAFTG 478					
b 377 LCPEBTLFSDGSPATGAPASSLGSSSTLQHAYAAFSPLLAMCYHJQCFQOKSTL 436					
Y 479 GHG-DALKVITGTFDSEPCSLCEKTRPVAHTYHRLORMPRFGQAOTrQPGVFGTM 537					
b 437 NPAYNITEVYGSAAANSPV-CSLQGQCPVCNTFLYFLRDPVQLGSRDPPVITGI 495					
Y 538 NSCYSDCDPLGRHYAPYLILRKPGDQZEEAATMDQTYRATERLFLDQEERLDRGACI 597					
RESUL 15					
T44001 major DNA binding protein [imported] - human herpesvirus 6					
C;Species: human herpesvirus 6					
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-JUL-2000					
C;Accession: T44001; T44188					
R;Iegawa, Y.; Nakano, K.; Kagaawa, M.; Chen, J.; Mori, Y.; Sunagawa, T.; Kawar					
J;Viroi, J.; 8053-8063, 1999					
A;Title: Comparison of the complete DNA sequences of human herpesvirus 6 variants A and B					
A;Reference number: 222732; MUID:9912319; PMID:10482544					
A;Accession: T44001					
A;Molecule type: DNA					
A;Residues: 1-1132 <168>					
A;Cross-references: EMBL:AB021506; NID:94995977; PID:BAAT8262-1; PMID:94996029					
A;Experimental source: strain HST; pop. variant B					
R;Domínguez, G.; Dambough, T.R.; Steamey, F.R.; Deburst, S.; Inoue, N.; Pellett, P.E.					
J;Viroi, J.; 8050-8052, 1999					
A;Title: Comparison of the complete DNA sequences of human herpesvirus 6 variants A and B					
A;Reference number: 222732; MUID:9912319; PMID:10482544					
A;Molecule type: DNA					
A;Residues: 1-1132 <168>					
A;Cross-references: EMBL:AB021506; NID:94995977; PID:BAAT8262-1; PMID:94996029					
A;Experimental source: strain HST; pop. variant B					
R;Domínguez, G.; Dambough, T.R.; Steamey, F.R.; Deburst, S.; Inoue, N.; Pellett, P.E.					
J;Viroi, J.; 8050-8052, 1999					
A;Title: Comparison of the complete DNA sequences of human herpesvirus 6 variants A and B					
A;Reference number: 222732; MUID:9912319; PMID:10482544					
A;Molecule type: DNA					
A;Residues: 1-1132 <168>					
A;Cross-references: EMBL:AB021506; NID:94995977; PID:BAAT8262-1; PMID:94996029					
A;Experimental source: strain HST; pop. variant B					
R;Domínguez, G.; Dambough, T.R.; Steamey, F.R.; Deburst, S.; Inoue, N.; Pellett, P.E.					
J;Viroi, J.; 8050-8052, 1999					
A;Title: Comparison of the complete DNA sequences of human herpesvirus 6 variants A and B					
A;Reference number: 222732; MUID:9912319; PMID:10482544					
A;Molecule type: DNA					
A;Residues: 1-1132 <168>					
A;Cross-references: EMBL:AB021506; NID:94995977; PID:BAAT8262-1; PMID:94996029					
A;Experimental source: strain HST; pop. variant B					
R;Domínguez, G.; Dambough, T.R.; Steamey, F.R.; Deburst, S.; Inoue, N.; Pellett, P.E.					
J;Viroi, J.; 8050-8052, 1999					
A;Title: Comparison of the complete DNA sequences of human herpesvirus 6 variants A and B					
A;Reference number: 222732; MUID:9912319; PMID:10482544					
A;Molecule type: DNA					
A;Residues: 1-1132 <168>					
A;Cross-references: EMBL:AB021506; NID:94995977; PID:BAAT8262-1; PMID:94996029					
A;Experimental source: strain HST; pop. variant B					
R;Domínguez, G.; Dambough, T.R.; Steamey, F.R.; Deburst, S.; Inoue, N.; Pellett, P.E.					
J;Viroi, J.; 8050-8052, 1999					
A;Title: Comparison of the complete DNA sequences of human herpesvirus 6 variants A and B					
A;Reference number: 222732; MUID:9912319; PMID:10482544					
A;Molecule type: DNA					
A;Residues: 1-1132 <168>					
A;Cross-references: EMBL:AB021506; NID:94995977; PID:BAAT8262-1; PMID:94996029					
A;Experimental source: strain HST; pop. variant B					
R;Domínguez, G.; Dambough, T.R.; Steamey, F.R.; Deburst, S.; Inoue, N.; Pellett, P.E.					
J;Viroi, J.; 8050-8052, 1999					
A;Title: Comparison of the complete DNA sequences of human herpesvirus 6 variants A and B					
A;Reference number: 222732; MUID:9912319; PMID:10482544					
A;Molecule type: DNA					
A;Residues: 1-1132 <168>					
A;Cross-references: EMBL:AB021506; NID:94995977; PID:BAAT8262-1; PMID:94996029					
A;Experimental source: strain HST; pop. variant B					
R;Domínguez, G.; Dambough, T.R.; Steamey, F.R.; Deburst, S.; Inoue, N.; Pellett, P.E.					
J;Viroi, J.; 8050-8052, 1999					
A;Title: Comparison of the complete DNA sequences of human herpesvirus 6 variants A and B					
A;Reference number: 222732; MUID:9912319; PMID:10482544					
A;Molecule type: DNA					
A;Residues: 1-1132 <168>					
A;Cross-references: EMBL:AB021506; NID:94995977; PID:BAAT8262-1; PMID:94996029					
A;Experimental source: strain HST; pop. variant B					
R;Domínguez, G.; Dambough, T.R.; Steamey, F.R.; Deburst, S.; Inoue, N.; Pellett, P.E.					
J;Viroi, J.; 8050-8052, 1999					
A;Title: Comparison of the complete DNA sequences of human herpesvirus 6 variants A and B					
A;Reference number: 222732; MUID:9912319; PMID:10482544					
A;Molecule type: DNA					
A;Residues: 1-1132 <168>					
A;Cross-references: EMBL:AB021506; NID:94995977; PID:BAAT8262-1; PMID:94996029					
A;Experimental source: strain HST; pop. variant B					
R;Domínguez, G.; Dambough, T.R.; Steamey, F.R.; Deburst, S.; Inoue, N.; Pellett, P.E.					
J;Viroi, J.; 8050-8052, 1999					
A;Title: Comparison of the complete DNA sequences of human herpesvirus 6 variants A and B					
A;Reference number: 222732; MUID:9912319; PMID:10482544					

C:Genetics:
A:Gene: 041
C:Superfamily: herpesvirus DNA-binding protein

Query Match	7.8%	Score 490.5; DB 2;	Length 1132;	Db 842 PDCYFHANKPLSALGFHDVTSGLTSKLVTONPVLPEVVLDPQPKFSSIQEYLVYVK 901
Best Local Similarity	22.0%	Pred. No. 1.9e-28;	Db 935 DVTSAADIEFOAKALLETEENPELWTAFTSTHLVRAMMQRPMVNLGJSISKYHAG 994	
Matches	278;	Conservative 180; Mismatches 535;	Db 902 LVLDG--VPNPIASLSKEPN---FGTFTS-----RSLVTEGTLKTFVSLA 944	
Qy	2 ENTQXTVTVPTGPGVYACRVEDIDLEBISPLAARSTDSDLALLPIMRNITVEKPTSS 61	Db 995 NNRVFOAQ--NWSGLNG-GKVNCPD----FPFDTRPFITACPRGGFICPVTPGPSSGN 1045		
Db	3 DENETVUSAPVSTAAWIYPPKEXELLDLVSILMERSPVVSPMLNMUTVENDSTT 62	Db 945 NREYFQFGQLGWIGSGVDRNLNPFTSSALQDFEMR-QKTIAATKFSEVIV----KKVR 998		
Qy	62 LAVVSGARTTGLAGAGITKLITSHFYPSSVFPHGKHVL-PSSAANPULTRACNAARERF 120	Db 1046 RETIISDQ-VRGIVS----GGAMVQLAYATVVRAGARAQHNNAFDWLSLTDD-E 1096		
Db	63 V-----KIPINNEGGILTITKITSFMPVCFFPHGEOLVGAEDGIDLIRICEQTRQKF 115	Db 999 RETIMEDFTEVVKGRVLSIVENLTNDIDPELLIAEVMR--DREDKPTMDMLFFVDGRE 1055		
Qy	121 GFSRQCGPPVDFGAVETTGAEICTRIGLEPENTILYLVITALKEAVPMCNVFLHYGOLDI 180	Qy 1097 FLARD---DEELEDQIQTQLETPWTVEGALAEV-----KILDEKITTAGDG---		
Db	116 HLQSFEVPTARKVTDIKA-LCSAVGKDADSVCVCHVACGNGFXELLIF-----AGLLI 165	Db 1056 ALAASIMLKLNHLVDMNVRDFFSIA-NLQSVFRAVSSNDAPVYDFSBILAEEDDQASGVLK 1114		
Qy	181 -----VHINHGDVIRIPLPVLPPQVLNFMDVNLPDPFNTHRSIGEGFYPTPFYNTG 2312	Qy 1139 --ETPT 1142		
Db	166 PCVBEQIQVQGVYESCVKIPLYSATLFETE-----ETISLSSTEFQERGFPLPA 216	Db 1115 CDEFET 1120		
Qy	233 LCHLIHDCTVAPAVALRVNVAVARGAAHAFDENHEGAVLPDITYTYFQOSSSSGTT 292	Search completed: January 30, 2004, 13:14:42		
Db	2117 LSBTLFYVFTSGTTLRFSENNTKELIDAGLKFTQDGEQTVKLAHKTYL-----266	Job time : 32 secs		
Qy	293 TARGARRNDVNSTSKPSGGFERRLASTMAATTALEAENVN-----TCIYEETPTDI 3446			
Db	267 ---GISGQKISAYEKD-----FLMLVDSVV-TELSSRSHAEYLDSDVYD-PSQI 309			
Qy	347 ---KEWPWFIGMGTPLPRNLALGSYATARVAGVIGAMFSPNSALYLTVE-----D 394			
Db	310 MNNDWPJIRNSNTTHAERMAQLNKGHLSSHLVIAFAPSILYSSKAFIPTNKQAFN 369			
Qy	395 SGHTEAKDGGPGSENRFYOFAGPHALANPQTDRDH-VLSSQSTSNSNTFSDYLLAI 453			
Db	370 SYNTQ-----ELLRSLSFRNALSSLTEDYNDNRKILKCDTSGKDJKFSANHLAYA 422			
Qy	454 CGGAPILLBLFYLERCDAAGFTGHGDALKYVGTGFDSE-----PCSLCEKH 503			
Db	423 CATCPOQIISYVWNLNMSV-----YNAENAYTIIYNHLVNCANSANLCEPCDGK 470			
Qy	504 TRVPCAHTTVHLRQLRNPREGATRQPIGVETMSQYSDCDPLGENYAPYLILRKPGDOT 563			
Db	471 CCCSCIGTAMYRVTGTRPAIPAVKKEPLVMSMERYIAEDILSFG-----KPKVSEL 525			
Qy	564 EAKRATMDTTRATLERFLDIEQERLLDRA-----PCSSSEGSSSYVDDHPTFR--- 613			
Db	526 KEIGKDQONTL-----LDRGKFKVSJFDYCKNSLIDPTGETENYRS 570			
Qy	614 -----RILDTRLRARIEOTTQFMKVLUETRDYKI--REGLESATHSMALTFDPYSGACP1 667			
Db	571 KKDFFVSIHGLTQCIBCVS---RCIVEMRTOQPPEQINCLQSFENVDTTPYATAFSF 627			
Qy	668 TNFLVKRTHLAVYQDLAISQCHCVPYQGOVQEGRNFRNQFQPVLLRFVDSL-----N 719			
Db	628 LTAFAVVKVLTQNL-----IVASGHVYDTPCTGNSISKWLQYQSYLGTPHSSYIK 683			
Qy	720 GGFIISTRITVLSSEGPDVSAPNPFLSQ--DA--PAGR---TFPGDILARSVEVIRD1R 770			
Db	684 KGFLNTRTVKV-----ASNVDMEQILDCDLYSGSKYKTTIOAKLCLSLMCILRFR 735			
Qy	771 VQNRRVVSQGNTNLSEAARASLVLGLASAYORQEKRVDMLHAGFLLKOFHGLJFPROMP 830			
Db	736 IKRPFENKSSKTAHHNPFKGN-----KHKKNPLSGCISLFLXHDKP-----781			
Qy	831 PNSK-SPNQNPTWLLQRONQPAKLTHEITT--IAAVKRFTERYAATFINLPTCIGE 888			
Db	782 PNVKISCLEMORPLANNVPTKTLIGNPEEKTFLKAFASITNTYDEIDIQPECUST 841			
Qy	889 LAQFMYANLILKYCDHSQYL---INTLTS-LITG-----ARRPDRFS--SVLHWIRK 934			

GenCore version 5.1.6
(c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

run on: January 30, 2004, 13:06:32 ; Search time 47 Seconds

(without alignments)

4062/725 Million cell updates/sec

Title: US-09-769-699-2

perfect score: 6294

sequence: 1 MENTOKTVTVPGLGVYAA.....DEFLDSGPIKHNITMEM 1203

scoring table: BL0SUM62 Gapop 10.0 , Gapext 0.5

searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

atabase :

A_Geneseq_19Jun03 : *

1: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*

2: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*

3: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*

4: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*

5: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*

6: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*

7: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*

8: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*

9: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*

10: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*

11: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*

12: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*

13: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*

14: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*

15: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*

16: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*

17: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*

18: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*

19: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*

20: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*

21: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*

22: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

23: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*

24: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*

Protein encoded by Streptomyces virid Mycelia sterilia C Enterococcus faeca Drosophila melanog Mechanical stress Human ORF1700 S. roseosporus dap Mechanical stress Rat osteoclast pro Rat osteoclast pro Rat protein 108. Rat CP. Rattus r Novel human diagno Novel human diagno Ge protein fragmen Streptomyces venezuelae vep Drosophila melanog Mycobacterium tuberculosis Human p53 target m Angiogenesis assoc Human peroxidasin Cephalosporin C ac Streptomyces venezuelae vep

ALIGNMENTS

RESULT 1

AAM52633

standard, Protein: 1203 AA.

ID AAM52633

XX

AC AAM52633:

XX

DE 18-FEB-2002 (first entry)

XX Varicella-Zoster virus (VZV) ORF29P protein.

XX ORF29P; open reading frame 29p protein; Varicella-Zoster virus; VZV; drug delivery; gene delivery; fusion protein; protein secretion.

XX Human herpesvirus 3 .

XX WO200154709-A1.

XX 02-AUG-2001.

XX PP 25-JAN-2001; 2001WO-02500.

XX PR 25-JAN-2000; 2000US-017791.

XX PA (UFCO) UNIV COLUMBIA NEW YORK.

XX PI Universitaet, S., Annunziato, P,

XX DR WPI; 2001-488744/53.

XX N-PSDB; AB02175.

PT Composition of matter for delivering desired compounds into eukaryotic cell, comprises Varicella-Zoster virus Open reading frame 29p protein

SUMMARIES

result No.	Query Score	Match Length	DB ID	Description
1	6294	100.0%	1203	Varicella-Zoster virus
2	3054	48.5	1452	Herpesvirus ICP8-G
3	3054	48.5	1452	HSV simplex vir
4	3016.5	47.9	1196	HSV-2 strain SB5 C
5	1566.5	24.9	623	HSV-2 strain SB5 C
6	1380.5	21.9	543	Macaque mulatta rba
7	620	9.9	1132	Varicella zoster V
8	183	2.9	35	Varicella zoster V
9	183	2.9	35	Varicella zoster V

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

09/769,699
2-21-04
Search Notes

CX	WPI; 2002-546283/58.	Db	649 BAHAMSLTUDPYAGCPCLLQLGRRSHLAVYDLDLASSQCHYTAGSVESRNPQFQ
DR	N-PSDB; AAD41974.	Qy	708 PVLRREFDVFDNGHISTRITVLSEG-PVSAENPTIGQDAPGRTEGDARVSVEVL 766
CX	Novel fusion protein comprising an ICP8 herpes viral protein and a detectable protein, useful for screening virus resistant cells and for identifying antiviral agents or agents that block the viral protein expression.	Db	709 PVLRARVMDDNNGPLSAKTIVLVASEGAI CAPSINGQTAAESSEGDARVTLGFP 768
X	Claim 9; Column 21-28; 29pp; English.	Qy	767 RDIRYENRYVFSGNCTNLSEAFARARLYGLASA-KQRGEVDMDLHGALQFHGLLFP 826
X	The invention relates to a fusion protein which comprises an ICP8 herpesviral protein and a detectable protein e.g. green fluorescent protein (GFP). Fusion proteins of the invention are useful for screening virus resistant cells and for identifying antiviral agents or agents that block the viral protein expression, for assaying the efficiency of antiviral agents or immunological reagents, to screen for infection in cells to detect virus-resistant and/or virus-susceptible cells, and to identify antiviral drug targets, antiviral agents or compounds, and/or agents that block the viral protein expression. The present sequence is ICP8-GFP fusion protein.	Db	769 KELRYKSRVIFAGASANASHAKARVASIQA-QKPDGRDVDTLQGFLLOPHAAIFP 828
Q	Sequence 1452 AA;	Qy	827 RGMPNPKSPNPNQFWTLQRMQMPADKLTHEETTIAVKRTTEAANPTNLNPPTCI 886
V	Best Local Similarity 49.8%; Prod. No. 2.5e-288; Matches 609; Conservative 200; Mismatches 366; Gaps 17; Indels 48;	Db	829 NGKPGQSNQNPNQFWTLQRMQMPADKLTHEETTIAVKRTTEAANPTNLNPPTV 888
Y	1 MENTQKIVT---VTPGPQGVYY -ACRVEDDLIEEFLRAARSTDPSDLALLPLMLNIVT 55	Qy	887 CELAFYMANILILXCDISQYLNTLTSITGARRPROPSDHWIRKDVTAAIDIQTA 946
Y	2 1 METPKPTATIKVPPGPQGVYYARACPSEGIE--LALLSARSQDAVAVAPLVVGLTVE 58	Db	889 SELANYMMANQILRYCDSTYF NTLLTAAIAGSRRPPSVQAAAW--SAQGAGLREGA 945
Y	56 KTFPSLSSLAVSGARTTGAGGITLKLTTSHFYPSPVYEGGGKHTLPSSAAPPNTBACNA 115	Qy	947 KALLEKTENPELMUTTAFSTHILVRAAMNQPRPVVLGIISYCHGAGANNRFOAGNSWG 1006
Y	59 SGPEANVAVVGSRTTGLGIIASLQKTPSHYSSSYVFGERHLIDESTOAPNLTRCLC 118	Db	946 RALDAVDAHPGANTSMRASCNLIPMARPVNLGIVSISYKGMGNDRFQAGRAWAS 1005
Y	116 ARERFGESRCGPPVDPGAVETTGAECITRIGLEPENTILLYVTAFLKEAVMCNFVFLH 175	Qy	1007 LNGGENVCPLFPTDRTFPLACPRGGPCTCPYGPSSCNRETTLSDPYRGIVTVSGGAMV 1066
b	119 ARERFGESDTPYTPGDKLKHETGEALCERGLDIDPRLLYVTVTGFKEAVCINNTFLH 178	Db	1006 LMGGKNAACPPLIFDRTRKFLVCPRAVGCAASNLSGAHESSLCEQRGISEGGAVAA 1065
Y	176 CGLDLVIWHGHDYTRIPFLPVOLMPDNRLVPDPFTTHASIGGSEPVYPFPTVGLCH 235	Qy	1067 LAYATVVRAGARAQHNAEDWISLTDEFEELARDLFLPHDOLIQUEPTWEGAL-- 1123
b	179 GGSKRVTGGAETHRIPYPLQFMPFSRVIAEPWVHNSIGMNTYPLPFERPLNR 238	Db	1066 SSVVATVRSGLPRTQQIQTEDPTALIEDEYSEEKMLTAAERGENTSDAALEVA 1125
Y	236 LIHDPCVIALPVALRVENTAVARGAAHAFDENHEAVLPDITYTYFOSSSSCTTAR 295	Qy	1124 -EATKILDEKTTAGDGETPTNLAFNFD -SCPSHDDTSNVNUNISGSNNSNISGTVPLGKRP 1180
Y	239 LIPFBAVVGPAALARCVNDAYAARAHALDENHGAADPADIIFTAFAASQG -KTPR 296	Db	1126 HEALALSVOLGNAE-----YENFDGEFC-----DDNATPFGPFGAPRGAFAFKRA 1173
Y	296 GARRNDVNSTSKEPSGCFERRALSIMAAADLHALVIENTGIVETPTDIKEVPMFIG 355	Qy	1181 PEDDEFLDLSGIPKRGHNITME 1203
Y	297 GGR-----DGGGKGDAAGFEEQNLASSWAGDAALLESIVSNAVEDEPPTDLSAWLCEGQ 351	Db	1174 FEGDDPPFG-EGPPDKKKGDLTLLDM 1195
Y	356 EGTLPLRNALGTSYVARYAGVGMVSPNSNLYLIVEDSNTENKDGQGSPSFRFQF 415	RESLT 3	ABG73-090
Y	352 DPAARANAVGAYLARAGLVAMPVETNSNALHLYDDAGPAPDDEHSK-PSPTRFFLV 410	ID	ABG73090 Standard; Protein; 1452 AA.
Y	416 AGPHLLANPOTDRDGHV-----SSQSTSSESSNTFSDVYLALLCGFAPLLARLLFL 468	AC	ABG73-090;
Y	411 PETHVAAANPOVDRGHTYPPGEGRPTAALPVGQT-BFAGHFLAMCGSPALLMFL 469	XX	14-APR-2003 (first entry)
Y	469 ERCDAGAFTGGH-G-DALKYVGTGFTDSPIPCSLCEGHTRPICAHTYHRQMPFQQT 527	DB	Herpes simplex virus (HSV) ICP8-GFP fusion protein.
>	470 ERCDGGTIVGROEMDYFVYVADSNQTDVPCNLCTDTRHCVHTLMRARHPRFASA 529	XX	Herpesvirus; ICP8-GFP; green fluorescent protein; virucide; fusion protein; HSV.
>	528 ROPIGVFGTANSQYSDCDPGLERYAPYLILRKPGDCEAATMOTYRATLIERLIEDQ 587	XX	Herpes simplex virus.
>	530 RGAIGVFGTANSQYSDCDPGLERYAPYLILRKPGDCEAATMOTYRATLIERLIEDQ 588	PV	Synthetic.
>	588 ERLIDRGAPCSSEGLSSVIVDHDPTFARIILDLTARIBOTTQEMKVLTETDVKREGIS 647	PP	US2002151033-A1.
>	589 LQYDQAVPTANGRLEIITITREALTVVNVROYDREYOLMLVNGRNPFERDGG 648	XX	17-OCT-2002.
>	648 EATHENALTFDPYSGAFCPITNFLVKRTHLAVVQDIALSOSCHCVEYQQVEGRNRNQ 707	XX	XX
>		PI	Knipe DM, Taylor TJ, McNamee EE;
>		DR	WPI; 2003-182512/18.
>		DR	N-PSDB; ABY15344.
>		PT	Novel fusion protein useful for detecting the presence, absence or amount of viral DNA replication and infection and as marker, comprises

viral protein derived from virus and detectable protein Disclosure; Fig 7; 30pp; English.

卷之三

The invention relates to a fusion protein comprising a viral protein with at least one function and a detectable protein, fused to maintain the open reading frame. The fusion protein is useful for determining whether a cell is a virus-resistance cell or a virus-susceptible cell, by contacting a virus that expresses the fusion protein with the cell to be tested, under conditions sufficient for the virus to infect the cell, and detecting the presence or absence of the fusion protein, where the absence of the fusion protein identifies a virus-resistant cell and the presence of the fusion protein identifies a virus-susceptible cell. The fusion protein is useful for identifying an anti-viral agent or an agent that blocks the expression of the fusion protein, by contracting a virus that expresses the protein, a host cell and the agent to be tested, in conditions sufficient to allow for the virus to infect the cell, and detecting the amount of the virus present, where a decrease in the amount of virus present identifies the agent. This sequence represents an ICP8-Green Fluorescent Protein (ICP8- GFP) fusion protein of the invention.

Sequence 1452 AA; $\Delta G = -10.4$ kcal/mol

Very Match 48.5% / Score 3054; DB 24; Length 1452;

Gaps 17; Indels 48; Mismatches 366; Insertions 609; Conservative 2001; Cachees 609; Files no. 235-268;

1 MENTQTKVT --- YPTGPIGYVY --ACRVEDDLEISFLAARSIDSDLLALLPLMRNLTVB 55

56 KTFPTSSLAVVSGARITGLAGAGITLKTTSHEYPSVEFEGGKAVTPSSAADPNTBZONA 115
1 MEIKRFLAINEFSPFLIVIAKALPSEGEL--LALLSARSEDADVAPLWGGEVME 58

59 SGPEANAVVNGSRTTGLGGTAVSLKTPSHYSSSVYFVGGRHLDPSQTAPNLTRICER 118

116. ARERFGFSRCQGPPVDAVETTGAECITPGLLEPENTILLYVTLFKEAVFMCAVTELY 175

11.9 ARRHFQFSDYTPREGDLKHHETTGAEALCRIGLDPDRALLVYVTEGEKEAVCINNTFLH 178

236 LIHDCVIAPMAVALVRNVTAVARGAHLAEDENHEGAYLPPDITYTYFOSSSSCTTTPR 295
237 JGDCAVIIISCPVWVHPLVQLFMEDSKVLLPFFNANHSIGENTYLPPLFFKRN 238

239 LLFEAVVGPAAVALCRNVDVARAAAHLAFDENHEGAALPADITETAPEASQG - KTPR 296

296 GARRNDVNSTS KPS PSSGG FERR LASIMA ADTAL HAEVIFTG IYEBPTD IKEWEMIGM 355

297 GGR-----DGGKGPGAGGFEQRLLASWMAAGDAALALESIYNAVEDEPPTDISAWPLCEQQ 351

416 AGPHLAANPQTDRGHVL-----SSOSTGSSNTTEFSYDYLICGGAPITARIVFV
417 DIAKXAVAGALLAKRAGLGVAFPSINSALHLIEVDLGSADPKUDHSK-PSYRFLV 418

4.11 PGTHVAANPQYDREGHVVPGFEGRTPALVGGTQ-EPAGGSHLAMLCGFSPALLAKMNFYL 469

4.6.9 ERCDAGAFTGGHG-DALKVTTGTFDSEIPCSLCEKHTRPVCAHTTVHRLRQNMFRQQAT 527

470 ERCDGGIVYGRQEMDVFRVYADDSNQTDVPCNLCTFDTTHACVHTTLMLRARRPKFASAA 529

588 ERLLDRGAPCSSEGSSVIVDHPPTERRILLDTLARIEOTTOMKVLYETRYDKIREGS 647
589 KESTAATWAGUNSMISUDUGLENHAASALKR - ADGSEYARIIMQETYRAATERVAELLET 588

589 LQYVDAQPTAMGRLETIITNREALHTVVNNVRQVVDREVEQMLRNLVTEGRNFKFRDGLG 648

648 EATHSMALTFDPYSGAFCPITNLYVKRTHLAVVQDIALSQCHCVFYGOQVEGRNFRNQFQ 707

卷之三

PT and treatment of infection or inducing immunological response in
mammal
XX Claim 10; Page 83-84; 748pp; English.

This sequence represents a Herpes simplex virus type-2 (HSV-2) protein sequence of the invention. This sequence was isolated from a HSV-2 strain SB5 (deposited as ATCC VR-2546) DNA fragment designated Contig ID 12. The proteins can be used for the treatment or prevention of disease, to induce an immunological response in a mammal or to identify inhibitors, activators or novel antivirals. Antagonists of the proteins can be used to inhibit a viral polypeptide. The DNA sequence or a vector containing it can also be used to induce an immunological response in a mammal.

Sequence 1196 AA;

Query Match 47.9%; Score 3016.5; DB 19; Length 1196;
Best Local Similarity 49.1%; Pred. No. 8.4e-285; Mismatches 211; Indels 41; Gaps 16; Matches 596; Conservative 367; Gaps 41;

b 4 TQKTYVTPGRLGYY--ACRIVEDDLEETSELALARSTDSDLALLPLMRNLTVBKTFTSS 61
7 TTTTKVPPDPGMITYGRA-PAGIEL- LSLIARSQADVAPlVLGTVSEGEAN 64
Y 62 LAVVSGARTGLAGAGITKLTTSHFYPSPYEVPGGGKHVLPSSAAPNLTRACNAARERFG 121
65 VAAVGSRSITGLGGTAVSLKLMRPHYSRSPSYVFFGRHLAPSTOAPNLTRCERARRHFG 124
b 122 FSRQCGPPVGDGAETTGAECLTRGPLEPNTILVNLVTAFLKEAVFMGNVFLHYGGDIV 181
b 125 FSDYAPRCPDLKHEETGDAICERGLQDPRALLYVITEGVREACVLSNTFLHGGNDKV 184
Y 182 HINQGDVIRIPLPVQLNFMDVNLPDPDNTHRSISIGGFYPTPPNTGLCLHLHDVC 241
b 185 TIGDAEVHRIPVYIQLQMFNPDFSSVIAFPNCNHSRSIGENFNPYLPFNNRPLARLFEAV 244
Y 242 IAPNAVAEVVRNTAVARGAAHLFDENHEGAVLPDIDTYTYFQSSSGTTARGARRND 301
b 245 VGPAAVALPARNTAVARAANHLFDENHEGALPADITTAPEASUG- KPQRGAR--- 299
Y 302 VNSTSXPSPSSGGFRRPLASIMAADTALEADEVINTGYYEETPFDIKEWPMIGMGTLP 361
b 300 - DGNKGPAFFQRLASVMAAGAALAESIVSMAVEDDEPPDTITMPWLLGQETPAA 357
Y 362 LNAIGSYTARVAGYIGAMYFSPNISALYLTEVEDSGMTAEKODGPGPSFRNFYQFAGPHLA 421
b 358 AGAGAYAARAAGVGAATFSTNSIALHTEVDAGADPKDKSK- PSFYRFVLPOTHVA 416
Y 422 ANPOTDRGCHVL -----SSQTSQGSNTESVYDYLALICGGPAPLARLLFYLERCDAG 474
b 417 ANPQLDRGCHVYVPGYEGRPTAPLPGGTO-EFAGEHLMICGGFSPALLAKMLFYLERCDGG 475
Y 475 AFTFGHG-DALKVTGTDSEPCSLCSKHTPVCATTHTVRLQRMFQATOPIQIGV 533
b 476 VIVGROENDVPRVYADSQTDVFCNLCTFETHACHTTLQRLRHPKFASAAGAIGV 535
Y 534 FGTMNSQYSDCDPLGNYAPYLIRKPSQTEAKATMDQTYRATERLFLDQEQRLLDR 593
b 536 FGTMNSAISDCDVLGNTAAFSALKR-ADGSENTRIMQETRASTERVMALEAQYVDQ 594
Y 594 GAPCSSEGLSSVYDHPFERRILDTLBARIEOTTQFMKVLYTETDYKIREGLSATHSM 653
b 595 AVTALGRLETTIGTREALHTVNNNIQLVDREREVOLMRNIEGFRNKFDGLAZANHAM 654
Y 654 ALIDDPYSGAFCPTNELVRRHAYQDLAUSQCHCVYQQVVERNPNOFOVLLRR 713
b 655 SLSLDPTCGPCLQLLARRNLAYQDALSQCHVFGASVERNPNOFOVLLRR 714
Y 714 FVDLNGGFIISTRSTVTLSEG-PSA5APNPTGQDAPAGRTFDDGLARYSVETDIRYK 772
b 715 VNDLNGFISAKTIVTYALSECIAIICAPSITAGQTAPESSPEGIVAVRLGFPELRYK 774
Y 773 NRVFSSNCNTNLSEAAARARLYGLASAYQROEKRVDMHGALGFLIKQFHGFLPRGMPPN 832

PT	Db	775 SKVLFAGASANASEAAKARYVASLQSYOKEDKRDVILLGFLGKPPAVIFPNGKPPG 834
PS	QY	833 SKSNPQWPHLQLORNQMPAKLTHEEITTAIAVKPFTEYAAINFNEPPTCIGELAQF 892
XX	Db	835 SNOQNPOWENTALQRNPLPARLSREDIEATAFKRFSLDYGA-INFNLAPPNSELLMY 894
JC	QY	893 YMANLILKCYCQASQYLIINTLTSITGARRDPSSVLAHTRKDVTSAADEIQARAKLKK 952
JC	Db	895 YMANOLIYCYCQHSTFINTLAVIAGSRRPVGQAAAAPP- -GGCAGLEAGRALMDS 951
JC	QY	953 TENVLPDWTATFSTHLVRAMMNORPMVYIGISISKYHGAGNNVRFOAGNWGLNGRN 1012
JC	Db	952 LDAHPGAWTMSMASCNLLRPMAAPMVNLGS SKYGRAGNDRFOAGNWASLIGRN 1011
XX	QY	1013 VCPFLFEDTRTRTFITACPRGCFICVYTGPSSGRNRETTISDQVQGILIVSGAMYQLAIAT 1072
XX	Db	1012 ACPLLFDFTRKFVACPRAGFVCAASSGGAAEHSLDEQRLRIIAEGGAAVASSVVA 1071
Y	QY	1073 VYRAVGJARAQHMAQDDWLSTTDDDEFARLDEBLHDQIQTQLETPWTVGAL---EAVKI 1128
Y	Db	1072 TVKSLOPRTOOLQEDWLAILEDYLSEMMEFTRALERGHGEWSSTDAALEYAHEAAL 1131
Y	QY	1129 IDEKTTAGDSETPTLNAPNFDSCBPSHDDTSNVNISGNSNISSTVPGKLRPBDDELD 1188
Y	Db	1132 VSOLGAGE-----VFNFGDFGDEDHAASTFGCLAA-AAGAGVARKRAFHGDDPG 1182
Y	QY	1189 LGI5P1KRGHTMTME-1203
Y	Db	1183 -EGPPEKK-KLTLLDM-1195
Y	RESIDU	5
Y	AAW	2197
Y	AC	AAW/2197 standard; Protein; 623 AA.
Y	XX	XX
b	AC	AAW/2197;
b	XX	XX
b	DT	13-JAN-1999 (first entry)
b	XX	XX
b	DE	HSV-2 strain SB5 Contig ID 15 ORF#32 protein.
b	XX	XX
b	XX	HSV-2 strain SB5; immunological response induction therapy;
b	KW	antiviral identification; viral protein inhibitor.
b	OS	Herpes simplex virus type 2.
b	XX	XX
b	PN	W0920016-A1;
b	XX	XX
b	PD	14-MAY-1998.
b	XX	PD-1997- 97W0-US20016.
b	XX	31-OCT-1997; 97W0-US20016.
b	PR	09-JUN-1997; 97W0-US20016.
b	PR	04-NOV-1996; 96US-0030279.
b	XX	(SMIK) SMITHKLINE BECHAM CORP.
b	PA	Chan JY, Dobrowski-Amaral CE, DeIvecchio AM, Dillon SB,
b	XX	Pi Easer KN, Kearny JJ;
b	PR	WPI: 1998-206847/25.
b	DR	Herpes simplex virus type-2 sequences useful in e.g. prevention and treatment of infection or inducing immunological response in mammal
b	XX	Claim 10; Page 123; 748pp; English.
b	CC	This sequence represents a Herpes simplex virus type-2 (HSV-2) protein sequence of the invention. This sequence was isolated from a HSV-2 strain SBS (deposited as ATCC VR-2546) DNA fragment designated Contig ID 15.
b	CC	Claim 10; Page 123; 748pp; English.
b	CC	This sequence represents a Herpes simplex virus type-2 (HSV-2) protein sequence of the invention. This sequence was isolated from a HSV-2 strain SBS (deposited as ATCC VR-2546) DNA fragment designated Contig ID 15.

CC Based on homology, this sequence is a major DNA-binding protein.
 CC The proteins can be used for the treatment or prevention of disease, to
 CC induce an immunological response in a mammal or to identify inhibitor,
 CC activators or novel antivirals. Antagonists of the proteins can be used
 CC to inhibit a viral polypeptide. The DNA sequence or a vector containing
 CC it can also be used to induce an immunological response in a mammal.

SQ Sequence 623 AA;

Query Match 24.9%; Score 1566.5; DB 19; Length 623;

Best Local Similarity 50.0%; Pred. No. 2.2e-143;
 Matches 313; Conservative 100; Mismatches 192; Indels 21; Gaps 9;

Qy 4 TQKTVTVPTPLGIVVY-ACRIVEDDLETSFLARSTDSDLALPIMRNLTVEKTFSSL 62

Db 7 TTTTYYKVPGPMDVYRCAPAEEL--ISLARSDDAVAPLVGLTVESGEANV 64

Qy 63 AVVSGARTTGLAGAGITKLTKTSHFYPSSVFPGGKHLVLPSSAAPNLTRACNAARERFGF 122

Db 65 AAVGSRITGLGGTAVSILKMPHYSPPSVVPHGRLLAPQAPNLTRCLBRARRHFGF 124

Qy 123 SRCGPPYDVGAVETTGAGICTRIGLEPENTYLIVVTALEKAVENMNVTEHYGGDIV 182

Db 125 SDVAPRPDCDLKHETTGDAJCERGLDPRALLYLVTEGPREAVCISNTFLHLGANDKV 184

Qy 183 INFQDVIRIPLIFPVQLMPDYNRLVPDPFENTHRSTGEGFTVTPPPYNTGLCHLTHDCVI 242

Db 185 IGDAEVHRIPVQLQMPMDFRVIAFPFNCRSGENENYPLPFNRPFLARLFEAVY 244

Qy 243 APMAVALVRVNNTAVARGAAHIAFDENHEGAVLPPDITYTYFOSSSGITTARGETARRNDV 302

Db 245 GPAAVALARNVNDAVAZAAAHHIAFDENHEGAALPAADTTFAPEASOG--KPGQGAR--- 298

Qy 303 NSTSKPSPGFFERRASIMAATDAHLAEEVFTNQGIEETPTDIKEWPMFIGMEGTLPRL 362

Db 299 -DAGNKSPAGGEFQRLLASVAGDAALESIVSMAVDEPPFDITWPLDEQETPPAARA 357

Qy 363 NAIGSYTARVAGVIGANTFSPNSIALYTEVEDSGMTEAKDGGPBPFSNRYQFAGPHLAA 422

Db 358 GAVGAYLARAGLIVGANFESTNSALHITTEVDAGAPDQHISK-DSFYRFFLVPPTHVA 416

Qy 423 NPOTDRDGHVLL-----SSQSTSNSNTESFYDYLALICFGAPLARIILFYLBRCDAGA 475

Db 417 NFDQLDREHVVGGYEGRPTAPLIVGGQ-EFGEHILAMLGCFSPALLAKMFLYLRCDGCV 475

Qy 476 FRRGHG-DALKYVTGTDSEIPCSLCCEKHTPVCANTHTVRLRQMPRTGQATQPQIGF 524

Db 476 IVGRQENDVFVYADSCQTDVNCNLCTFHCAHTTMLRLRHPKFASAARGAIGF 535

Qy 535 GTMNSQYSDCDPGLNTAPYLRLKPDQTEAKAMQDTYRATERLRLFIDLEQBRLLDRG 594

Db 536 GTMNSAYSTCDCVLGNYAAFSALKR-ADGSENTRTIMQETYRAATERVMAELAQYDOA 594

Qy 595 APOSSEGSISSVTDHPTPFLTLR 620

Db 595 VPTALGRLETTIGTREALHTVNNIK 620

Qy 892 FTMANLILKYCDHSQSYLINTTSITGARRPDPSSVHLWIKRDVTSAAIDIETOAKALLE 951

Db 241 YMANQOLRYCDHSITVPTLINTAVAGSRRPGVQAAAAPQ---GGAGLEGARALMD 297

Qy 952 KTNENLPELWTIAFTSTHLYRAAMNQRPMYTLGIISISKYGAAGNNRVOAGNHSGLNGK 1011

Db 298 SLDAHPCAWTMSFASCNLURPMARPMVTLGLISKTYGMAGNDRVQAGNWASLUGK 357

Qy 1012 NYCPLTEDPDRRFLIACPRGGFCIPVTPGFSGNRETJLSQVQGIIIVSGGAMYQLAYA 1.071

Db 358 HACPLLFEDRTRKFLIACPGAFVCAASSLGGGAHEHSLSCELRGTTIAEGGAIAVSSYFV 417

Qy 1072 TVURAGARIAQHMAFDDWLSLTDEFDLARDLEEHQDITQLETPWVEGAL---EAVK 1127

Db 418 ATVKSLGPRTOQEDWLALDEBLSMEMFTTRALERGHEWSSTDAALEYAHEAEA 477

Qy 1128 ILDEKTTAGDGETPINLARNFDSCBPSHPTTSNTLNISGSNTVPELKLKPBDDEL 1187

CX Herpes simplex virus type 2.

Db 478 LVSQQLGAGE-----VFNFGDFGDEDHAAASFGGGLAA--AGAAGVARKRATHGDFF 528
 2Y 1188 DISGIPKHNITMEM 1203
 Db 529 G-EGPDEKK-DLTDM 542

RESULT 7
 AAB53126 standard; Protein; 1132 AA.
 IC AAB53126;
 JT 28-FEB-2001 (first entry)
 OX Macaca mulatta rhadinovirus 17577 RRV ORF6 protein SEQ ID NO:3.
 OX Macaca mulatta rhadinovirus 17577; RRV; rhesus macaque rhadinovirus;
 CW genome; Kaposi's sarcoma-associated herpesvirus; KSHV; interleukin 6;
 CW IL-6; macrophage inflammatory protein; MIP; diagnosis; vaccine;
 cytotoxic; anti-HIV; gene therapy; infection; Kaposi's sarcoma;
 lymphoproliferative disorder; B-cell hyperplasia; lymphadenopathy;
 splenomegaly; hypergammaglobulinæmia; autoimmune haemolytic anaemia.
 IS Macaca mulatta rhadinovirus 17577.
 IN WO200028040-A2.
 X D 18-MAY-2000.
 X F 05-NOV-1999; 99W0-US26260.
 X R 06-NOV-1998; 98US-017507.
 X R 20-NOV-1998; 98US-019409.
 X A (UYOR-) UNIV OREGON HEALTH SCI.
 X I Wong SW, Axthelm MK, Searies RP;
 R WPT; 2000-375552/32.

New rhesus rhadino virus for producing non-human primate model useful for testing potential treatments and efficacy of the candidate vaccine for conditions associated with RRV infection. -
 S Claim 5; Page 123; 14pp; English.

The present invention describes a novel rhesus macaque rhadinovirus called macaca mulatta hadinovirus 17577 (RRV). AAC64754 represents the RRV genome sequence, and AB53123 to AB53204 represent the proteins encoded by the genome sequence. The present invention also specifically claims the individual open reading frame (ORF) nucleotide sequences from the genome which encode the individual proteins, but these sequences are not given. A non-human animal infected with RRV can be used for testing the efficacy of drug in the treatment of condition associated with infection with RRV such as Kaposi's sarcoma, lymphoproliferative disorders, B-cell hyperplasia, lymphadenopathy, splenomegaly, hypergammaglobulinæmia or autoimmune haemolytic anaemia, by administering the drug to a immunocompromised non-primate, preferably rhesus macaque monkey obtained by as a result of infection by Simian Immunodeficiency Virus (STV). RRV is useful for producing non-human primate model for testing potential treatments for conditions associated with RRV infection. It is also useful for testing the efficacy of the candidate vaccine against RRV infection or conditions associated with its infection by administering the vaccine to the subject capable of infection with RRV, inoculating the subject with RRV and observing the effect of vaccine. AAC64755 to AAC6765 and AB53205 to AB53213 represent sequence used in the exemplification of the present invention.

Sequence 1132 AA.

Query Match 9.9%; Score 620; DB 21; Length 1132;

Best Local Similarity 22.6%; Fred. No. 2.7e-50;
 Matches 290; Conservative 212; Mismatches 506; Indels 274; Gaps 50;
 Qy 1 MENTQKTVTPTGPLGYVACRVEDDLEESIFLAARSTDSDLALLPLMRNLTEKTFPS 60
 Db 11 LDNQGS-RATIGACCXYVAKSKQDPPAASILGNRPSGSGVSLPILYQGYHEFPL 69
 Qy 61 SL-AVYSGARTGTLAGAGITKLTSHFYSSVEFHGGKHLVPLSSAAPNLTIRGNAARER 119
 Db 70 TVKAAYKKVDDTTLA----VKVTC--PHREVIYTHNASLFRPVFDGCLINELGEARAL 122
 Qy 120 FEFSSRC--QGP--VDGAVETGABICTRIGGLEPNTILYLWVTALPKS--AV 166
 Db 123 PGYTOFIEPGPHSIWNPL--CPOLP-DDEMEGUVVTEGFERLWRGCLVPV 175
 Qy 167 FMCNVFLHYGGDLDTVNHNSDPIVIRPLFPVQLFMDVNRVLPDPFENTHRSIGEGFYPT 226
 Db 176 EQTQQ----VQIAQRQAFKPVLPYDDELPAFHGRM- 207
 Qy 227 PFTYNTGUCHLHDVCYAPMAVALVRVNTAVARGAAHLADENHEGAVLPPDITYTYFQS 286
 Db 208 RFYHDKDSVAYLDSLFTSIAAQALPLKDVTAVIATEXOFMDQHNYKIAKI--YQAKOF-- 262
 Qy 287 SSSGTTTARGARRNDVNNTSKPSGGPFRLSISMAADTALEAEVIFNTGIVEETPPDI 346
 Db 263 -----STLPLXTTDGSSHMIVDSVVAELASYGCMF---LECPQDA 300
 Qy 347 KE----WMFPIGMEGTIPRLNALGTSSTARVAGVIGAMVFSPIPSALSYLTVEDSGMTEA 400
 Db 301 CELLYNDSNPIDGCDSPARVANLERSAQPVHAGQLPRAINSVLYLKTKQKQAPRGQ 360
 Qy 401 RDGGGPSPSFNRYQFAGPHLAAFPQTDRGK----VLSLOSSGSSNTESFSDYDYLALJC 455
 Db 361 K---GDUVNDNSFFLQHGQFLNEATIENGSEAFPKGVSNAUDGSSFPY--HLAYAAS 415
 Qy 456 FGAPIALARLFYLERCDAGAFTGGHDALKYMTGTDESIPPSLCEKHTRPICAHTVHR 515
 Db 416 FSPHLLAKUQYMFQLOHEHKSTSNTQAFNMVHVTGTAANSEMTCIHCNTPATCLNLTFYR 475
 Qy 516 LRQMPBREQANTROPIGFQFTGTSNQSDCDLIGNYAPYLIRKPGQQT-EAKATMQDT 573
 Db 476 LQDFPAPVTTPORDPPYVWTGAGT-FNDLTLGNEFSFRDREEDGNPADEPKYTYWQ 534
 Qy 574 YRATERLFLDIBERBLDGRADPOSSBLS----SVVHDPPTPRRLDILTRARIEQT 626
 Db 535 CQTYKEK-----SAGCTTBDHNTYLTNIQSFILRVPKDSDTJDGE 578
 Qy 627 TTOPMKVLYBTDRYKIREGLSLAETHSMALTFPPYSGAFCPTTINFLYERTHLAVVODIALS 686
 Db 579 VMKETVNSMTK-NNNFPRHVKSYHIIQCCVWQOPCAVFLNLYKSLWMIQDCLP 637
 Qy 687 QCHCVYFGO-----OVEGINFRN-----OQPYVLRPPEVDLF 718
 Db 638 -YIMYI6DNNPANGLIPSEWWJRHCFOTLWTFKACLDRLAVLGEBLKVRHDWFCDFE 695
 Qy 719 ---NGCSF1SRTSITVLSSEGPSVSAAPNPLGQDAPGRTDGLARVSEVIRDIVKN 773
 Db 696 DTDAISGJNMAPKMQR-----TARAMMVYPSKIKN 729
 Qy 774 RYVFSGNCTNLSEARARLV--GLASAYQFQEKYDMLHSAALGFLLKQPHGLLPRGMP 830
 Db 730 RILFSN-TAGGEAVQGCFVKETGTRDY----VIVGPMFLNSHRALE---774
 Qy 831 PNSKSPNQWFTLQLQNQMPADK-LTHEBTITIA-VKRFTEETAAINFNLPTCIGE 888
 Db 775 PDKTKTAAALYWKISKQTNKTPWVKDYPDDELAELSYVKTNSLAETBNVLDVDDSLMS 834
 Qy 889 LAQYMANLILYKCDISQYLTNTLTSITGARRPDPSSVHLWIKDVTSAADETQATA 948
 Db 835 YARIKINGAILAACQGQIYATTL-HCLTPVLTQDABEYPH---VUGSAAAATPVAY 888
 Qy 949 LUB---KTNLPELWTAFTSTHLYVRAAMNQRPMTNGISISKYHAGCNRVOAGN--1003

Db 889 LAEIRGRT----ALTVQTTARQPVAATGRLRPVITVPMVNKKYTGYNNNNVFHCGNLG 943
 QY 1004 -WSGLNGKKNVCPFLF-FDRT-----RFRFIACPREGFCIVTGPSSGNETTISDQV 1054
 QY 944 YFAGRGVDNRLWPESSPFKXTGSAMLRGRHNMTP---IIDLILKEAAG-QTISTFEA 998
 Db 1055 RGIIYSGGAMQ----LAIVATVDRAVGARAQHNAFDWMSLTDDEFIARDLBELH 1106
 QY 999 ESRKRSVQAILEDKDNPNLKSVTLETRHLGKGCD----LSSEDOVYIGDYMCLT 1052
 QY 1107 DQIQTLE----TPWTEVGALPAVKILDEKTTAGDGETPTNLAFNFDSCBPHD--T 1157
 Db 1053 DEVLFDTNDIACGSYWPTE--DAGALIEDRQDADLQFVBDSDIATASCOPEBLT 1109
 QY 1158 TSVNLINISGSNISGSTVPLKX 1179
 Db 1110 PS-----AGALLAGKR 1121

RESULT 8
 AAY26935 standard; peptide; 35 AA.
 ID AAY26935;
 XX AC AAY26935;
 XX DT 21-DEC-1999 (first entry)
 DE Varicella zoster virus type 29 recombinase D35E motif.
 XX Recombinogenic motif; retrovirus; integrase; invertebrate; transposase; consensus; vaccine; gene therapy; allergy; autoimmune disease; cancer; cardiovascular disease; graft rejection; hematopoietic disorder; immunosuppressive disorder; immunoproliferative disease; jaundice; immunodeficiency disease; infectious disease; inflammatory disease; septic shock; metabolic defect.

XX Varicella zoster virus
 OS US5959074-A.
 XX PN 28-SEP-1999
 XX PD 28-FEB-1997/ 97US-0807332.
 XX PR 01-MAR-1996; 96US-0012616.
 XX PR 02-AUG-1996; 96US-0023054.
 XX PA (NAJB-) NAT JEWISH MEDICAL & RES CENT.
 XX PI Gelfand EW, Drayfus DH;
 XX DR WPI; 1999-561074/47.
 PT Peptides derived from a Herpes virus recombinogenic motif useful for the production of vaccines and in gene therapy -
 PT Claim 1; Column 37-38; 29pp; English.

XX The invention relates to novel recombinogenic motifs found in (retro)viral integrases which have similarity to invertebrate transposase molecules. The peptides, or their homologues, have the following identifying characteristics: (a) an initial and terminal amino acid comprising Asp or Glu, separated by 34-39 amino acids, especially 35 amino acids (the D35E motif); (b) a high probability ($p < 0.05$) of alignment with a consensus sequence (AAY26941), as determined by: (1) aligning the initial and terminal residues of the homologue sequence with those of the consensus; (2) aligning the intervening residues of the homologue with those of the consensus by maintaining the spacing of the consensus while, if necessary, altering (by the insertion of spaces or the deletion of residues) the spacing of the homologue; and (3) generating a distribution of 1500 random sequences of amino acids between the initial and terminal residues and identifying the percentage of aligned residues. The recombinogenic motifs may be useful in the

CC production of vaccines and in gene therapy to treat allergies, CC autoimmune diseases, cancer, cardiovascular diseases, graft rejection, CC hematopoietic disorders, immunosuppressive disorders, CC immunoproliferative diseases, immunodeficiency diseases, infectious CC diseases, inflammatory diseases, jaundice, septic shock and other CC immunological, genetic or metabolic defects.
 XX Sequence 35 AA;

RESULT 9
 AAB73931
 ID AAB73931 standard; peptide; 35 AA.
 XX AC AAB73931;
 XX DT 29-MAY-2001 (first entry)
 DE Varicella zoster VZV29 protein recombinogenic sequence.
 XX Varicella zoster virus; VZV29; herpesvirus; D35E; anti-viral; immunosuppressant; immunostimulant; cytostatic; anti-allergic; cardiant; anti-inflammation; gene therapy; retroviral integrase; herpes virus replication; V(D)J recombination; immunoglobulin class switching; allergy; autoimmune disease; cancer; cardiovascular disease; graft rejection; haematopoietic disorder; immunological disorder.
 XX KW
 XX OS Varicella-zoster virus.
 XX PN US6187584-B1.
 XX DR WPI; 2001-256131/26.
 PT New nucleic acid molecules encoding peptides with transposase activities, useful as immunosuppressants, anti-viral agents, or vectors for gene therapy of diseases, e.g. allergies, autoimmune diseases, cancers or graft rejection -
 PT Claim 1; Column 25; 31pp; English.
 PT The present sequence is provided in a specification relating to a new CC isolated nucleic acid molecule comprising a nucleic acid sequence CC encoding a peptide consisting of one or ten amino acid sequences CC defined in the specification or homologues of these sequences. CC the nucleic acid molecules are useful for producing peptides CC important in the regulation and function of herpes virus replication, CC V(D)J recombination, retroviral integrase function or immunoglobulin CC class switching. The isolated nucleic acid molecules and the peptides CC they encode are useful as immunosuppressants, anti-viral agents or CC vectors for gene therapy. These are particularly useful in protecting CC humans from a variety of diseases, e.g. allergies, autoimmune diseases, CC cancers, cardiovascular diseases, graft rejection, haematopoietic CC

WO200040704-A1.
13-JUL-2000.
06-JAN-2000; 20000WO-US000445.
06-JAN-1999; 99US-0115445.
05-FEB-1999; 99US-0118848.
05-JAN-2000; 20000US0477962.

(REGC) UNIV CALIFORNIA.
Shen B, Du L, Sanchez C, Chen M, Edwards DJ;
WPI; 2000-465974/40.
N-PSDB; ARA58471.

New bleomycin gene cluster components useful for peptide and/or polyketide metabolites, especially bleomycin, production and for chemically modifying biologics.

Disclosure: Page 97-137; 162pp; English
AAB07556-78 represent proteins encoded by open reading frames (ORFs) 8 to 30 of the BLM (Bleomycin) gene cluster. The proteins encoded by the gene cluster are useful for producing peptides and/or polyketide metabolites, especially bleomycin or bleomycin analogues. They are also useful for chemically modifying biological molecules to produce branched methyl groups, and for coupling amino acids and fatty acids. They may be re-arranged with an apo-carrier protein and coenzyme A to produce a holo-carrier protein. The BLM gene cluster or catalytic domains can be used individually or collectively to produce thiazoline, thiazole, thiazine, and bithiazoline-containing microbial metabolites. The BLM gene cluster may also be used to produce other metabolites.

Sequence	1841 AA;	Query Match 2.5%;	Score 156.5;	DB 21;	Length 1841;
		Best Local Similarity 19.6%;	Pred. No. 0.00015;		
		Matches 268;	Conservative 133;	Mismatches 470.	Truncat. 492.
					Score 156.5;
					Length 1841;

Qy	857 HEITTI-AAVKRFTTEY--AAINFINLPTCIGELAQFYMANLILKCDHSQYLINTLT 913	Db	11879 PSDXSPPPRTPRSGRRGRIAARRRACRPAPHSPTRXTGAAPGHARRCRXGRRGIPG 11938
Db	1377 HPERATVPAAKTPREPVIACTALDVEPLGLDE---RLADLIVELGAR--ETTV 1429	Qy	147 ----LEP-----ENTIYLVYTALEPKAEVEMCNVFLHYGG 177
Qy	914 SITGARR ---PRDPSSYLHWKRKDVTSAADETQAALKEXTENLPELTAAFTSTHL 969	Db	11939 ARGKAEPHMASRXSLRRSPDAPSNLLENNAKORISAKGKA----GG 11988
Db	1430 TACGRGRRRTPCP-----TROPLPA-----PERPAPVRPGVYL 1463	Qy	178 LDIVH-TINHD-----VTRIPLFPPVLFMPD-----V 203
Qy	970 VRAAMNQRPMKVNLGVISIXYRGAA-----GNRVRFOA-GNWGLNGKNTCPLFEDRT 1022	Db	11989 LSPSHGCCRHGQSHAVGORDALVORLPPVPU-LGPVAIGVGDAELLERHTGYAERABAGVA 12047
Db	1464 VCGLGIGG----IGLHLAYLGLGARTTVLTHRPAPGADGLPAGH--PEAAVRR 1515	Qy	204 NRYLPDPFNTTHRHSIGEGFVY-PTPFYNTGCHLHIDCVIAMAVALRVNTTAVARGAA 262
Qy	1023 RRFIIACPRGGFICPVTGCFSSGNBRETTLSQD-----VRGJIVSGG--- 1062	Db	12048 QRVAFVV----SLPVQYTLDVPAGQGVQLLRQFOIQSOLG--RARQYVDLPR---
Db	1516 LRSLAA-----TGATVVRVRAADLTDHDAMRALADEVEQAHGPVRGTYHAAQVPTD 1565	Qy	263 HLADENHE----GAVLP-PDTITYTFQSSSSGTTTARGARRDNVNSTSSPSGGFER 316
Qy	1063 -ANVOLAIYATVYTRAVGARAQHAFDWSLTLTDEFARDLDEL 1105	Db	12049 QPAFGDOPPAHPVGHVHPVPOVYAAIPVQRDAVAVQOPGGQNRNDL-----LR 12143
Db	1566 AgmIQRDRAGTDAALAKLTGTLL-----VLEDEVFAHRDLDFL 1603	Qy	317 RLAS--INAADTALHAEVTFNTGIVYBTPTDIREWPNFIGNEGTPLRNALGSYTTARVA 373
RESULT 11			
ABP76682	ABP76682 standard; Protein, 19938 AA.	Qy	12144 ELLGPVWVAAAPRHAHQ-----PMGAGV-GEGDHVAALGLRRVGRVG 12184
XX	DE Streptomyces viridochromogenes Avi gene cluster polypeptide frame 6.	Qy	374 G-----VIGANVESPNSA--LYLTEREDS--GNEAKDGPGPSFNR 411
AC	KW Avilamycin; antibacterial; virucide; protozoicide; fungicide; infection; infection; enzyme.	Db	12185 GDRRLVPGARDLDAVHLVNGDLHDPRAGQAHLECLDLHIGTYERRVGDRPVDVR 12244
XX	OS Streptomyces viridochromogenes.	Qy	412 FYQ-----FAGPHLAANPOTDRDGHVYSSQSTGSSNTPEVDYLALICFGAPILLARLL 465
XX	DE Streptomyces viridochromogenes Avi gene cluster polypeptide frame 6.	Db	12245 FGRBVDRHIVAGHHLGQPGGI-ADYAVHEGVPGDRLOVQPGGQVLYQHGHPERIAV- 12302
XX	KW Avilamycin; antibacterial; virucide; protozoocide; fungicide; infection; infection; enzyme.	Qy	466 FYLERCDAGAFIGGHHGDALKVYTG----TDFSETPCSLOCERKTRPVCAHHTYHLROR 519
XX	OS Streptomyces viridochromogenes.	Db	12303 -----PIQQFAHVYGADEARSYSTDQDVYSPSP--RTRAAVSHVPSRLHER 12344
OS	XX WO200268436-A1.	Qy	520 MFRFGQATRQPQGVFTGMNSQYSQDCDPLGNYTAPYLILRKKGQDQEAAKTMQDYRATUE 579
PN	XX PD 06-SEP-2002.	Db	12345 VPLLRRREVQOPTWT-RLWSPSTSRCAVDVRGAATFSVSRLTESTRAPSPIEAPRSTIE 12403
XX	XX PR 24-AUG-2001; 2001WO-EP09815.	Qy	580 RLFDLDEQRLL-----DGAPECSSEGLSLSVTDHPTTRR-----ILDTLR-ARIE 624
XX	XX PR 25-FEB-2001; 2001DE-1009166.	Db	12404 -YWLSCATRQIPLSMEVKGPKXESCTT-VSGPTAGPTIREPEIFAPASTTRPTSSLIA 12460
PA	(COMB-) COMBINATURE BIOPHARM AG.	Qy	625 QTTOFMKVLYVETDRYKIRECLSEATHMSA-----LFDPYSGAFCPITFLYKRTLH 677
XX	XX PI Weitnauer G, Muehlenweg A, Trefzer A, Bechthold A;	Db	12461 SSTSPSNRV--SRDSSRCRISRISVTLTSFQRSIVTEPPTS---CP----- 12503
XX	XX WP1; 2003-01-18650/01.	Qy	678 AVYQDIALSOQHCVFQGQQUEGRNNO-----POPVLRRRFVDFLNGGF 722
DR	DR N-PSDB; AB237516.	Db	12504 -----WSAHC--RASYTSSRQDRGMMPATASCTAELKAYPTIARSL--LGSGCF 12551
XX	XX PR 07-SEP-2002.	Qy	723 ISTRSITVTLSEGSPISAPNPIL-GDADAPAGRTPDCDLARVSEVTRDRVKNERVPSGNC 781
XX	XX PS Example 1; Page 68-301; 319pp; German.	Db	12552 -STTAVTRPFS---SSATPMLRGSATRASMTMCAGLSR-RKSATKSTMPTXLSPRYM 12606
XX	The invention relates to avilamycin derivatives (I) with antibacterial, treatment of infections (bacterial, viral, protozoal or fungal), in human or veterinary medicine, particularly where caused by <i>Staphylococcus aureus</i> . (I) are more hydrophilic than known avilamycins. The present sequence is that of an avilamycin synthesis enzyme from the Streptomyces viridochromogenes Avilamycin A biosynthetic gene cluster (AB237515-AB237516).	Qy	782 TNLSEARARVGLASAYQDQEKRVDMHGLGJLKKOPHGLJFPRGMP-----PNS 833
CC	CC New avilamycin derivatives, useful for treatment of infections, and nucleic acid encoding avilamycin synthesis enzymes	Db	12607 .KSSSPRKSRTSTASA-----SPRGASCKXVTSRPNs 12640
CC	CC Example 1; Page 68-301; 319pp; German.	Qy	834 RSPNPFWW-----TLLQRNONPA---DKLTHEITTAIAVTRFTEYALI-----NFNL 881
CC	CC The invention relates to avilamycin derivatives (I) with antibacterial, treatment of infections (bacterial, viral, protozoal or fungal), in human or veterinary medicine, particularly where caused by <i>Staphylococcus aureus</i> . (I) are more hydrophilic than known avilamycins. The present sequence is that of an avilamycin synthesis enzyme from the Streptomyces viridochromogenes Avilamycin A biosynthetic gene cluster (AB237515-AB237516).	Db	12641 -APSPTEAWIASAAVAPTPTIFSDDASLMESSTXERTVLLATOTSCLAEVCTGIRSLEPT 12699
CC	CC Sequence 19938 AA;	Qy	882 PPTCTIGELAQFYMANLILKCDHSQYLINTLTSITGARRPDPSSVLMWIKRDVTSAAAD 941
CC	CC PS	Db	12700 EPT-----NNNSAFISITLRS-----PSDOV-----VSELOQASN 12730
CC	CC The invention relates to avilamycin derivatives (I) with antibacterial, treatment of infections (bacterial, viral, protozoal or fungal), in human or veterinary medicine, particularly where caused by <i>Staphylococcus aureus</i> . (I) are more hydrophilic than known avilamycins. The present sequence is that of an avilamycin synthesis enzyme from the Streptomyces viridochromogenes Avilamycin A biosynthetic gene cluster (AB237515-AB237516).	Qy	942 IETOAKALLEKTEENIPEL-----WTAFTSPHVLRAAMNQR-PMVYLGISISKYHGA 992
CC	CC Best Local Similarity 19.4%; Pred. No. 10; Mismatches 129; Conservative 246;	Db	12731 VSAQDPRVQMFQVPRPGRLARPHPAVPCXHRLSQCVDAPWSAV-GCSVRR-SV 12787
CC	CC 102 PSSAAINLTRACNAAREREFESR-----CQGPBPVDGAVETG-----AEICTRLG----- 146	Qy	993 AGNNRPUQADMWSQMGSLNGGGRNCFLIFTFRTRFLIACPRGGF-COPVTPGSSGARETFLSD 1052
CC	CC 12788 RXSTMMSLMTGSSFORMLAGTS-----RPRP-----SSRRTTIVIT 12823	Db	

2Y 1053 QVRGILIVGGAMVQLAIYATVRAVGARIQHMAFPDDWLSLTDDBEFLARDLSELHDQIQT 1112
 Jb 12824 ESSPSTSG-----RETCRAYGASRISCA---CERTSSSS----- 12855
 Yy 1113 LETPNTVEGALEYKILDEKTTAGEDGETPTNLAFNEDSEPSHDTTENVLNISGSNSIGS 1172
 Jb 12856 -----TRERSSGD----- 12885
 Yy 1173 --TYPGLXRP 1180
 b 12886 AGCSTTGXRPP 12896

:RESULT 12
 ABB395B standard; Protein; 3210 AA.
 X ABB395B;
 C X
 X 29-MAY-2001 (first entry)
 X Mycelia sterilia cyclic depsipeptide synthase.
 X Mycelia sterilia; cyclic depsipeptide synthase; anthelmintic;
 X PR022.
 S Mycelia sterilia.
 X WO200118179-A1.
 X D 15-MAR-2001.
 F 07 SEP-2000; 2000WO-JP06103.
 X X 07 SEP-1999; 99JP-0253040.
 R 06 APR-2000; 2000JP-0104291.
 X X
 X (MEIJ) MEIJI SEIKA KAISHA LTD.
 I Midoh N, Okakura K, Miyamoto K, Watanabe M, Yanai K, Yasutake T;
 I Aihara S, Futamura T, Kleinkauf H, Murakami T;
 WPI; 2001-265970/27.
 R N-PSDB; AAF79702.
 X X
 T Novel cyclic depsipeptide synthase and gene encoding it for efficient
 T production of anthelmintic substance PF1022.
 X X
 Claim 1; Page 59-80; 92pp; Japanese.
 X X
 C The present sequence is a cyclic depsipeptide synthase from Mycelia
 C sterilia. The present protein, or a protein containing an addition,
 C deletion and/or substitution of one or more amino acid residues is useful
 C for the efficient production of the anthelmintic PF1022 (cyclo(D-lactyl-
 C L-N-methylleucyl-D-3-phenyllactyl-L-N-methylleucyl)-D-lactyl-L-N-
 C methylleucyl-D-3-phenyllactyl-L-N-methylleucyl)).
 X X
 Sequence 3210 AA;

Query Match 2.0%; Score 123.5; DB 22; Length 3210;
 Best Local Similarity 18.7%; Fred. No. 0.66; Mismatches 79; Indels 177; Gaps 24;
 Matches 111; Conservative 79; DR WPI: 1999-070095/06.
 DR N-PSDB; AAX20196.

Y 298 RRNDDNSSKPSGSPGGFERRASIMADTAHAEVFNTGYYEEPTD1KEWPMIGME- 356
 b 2550 RRAQIAPKSQAAPA-----KPKVQDPPFTNLEALL-----CEEEFAEVIGMVEY 2591
 Y 357 GTDPLRNALGASYTAVGIGANVFSNSALYLTVEEDSGMTAEKGPGPSFNRFYQFA 416
 b 2592 GVDNHFPQGGHSLLATKVALVSLRNGYSVRDQ----- 2630
 Y 417 GPHLAANPQTDRGHVLS-----QSTGSSNTEFSDVYALICFGAPPLRL 465

Db 2631 -PVISDLAVTLRQGLTLENAIPATPDGYWEQTMSAPTPSDMEAVLCKEPAADVLG-- 2686
 Qy 466 FYLERCDAGAFGTGGHG-----DALKYVTGTFDSETPCSLCEKTRPVCAHT 511
 Db 2687 VEVSATDSFFDGGGSMATKLAARISRRDVPSIKDIFDHSPVPLNARK----- 2737
 Qy 512 TVHRLRQMRMPRGQATROPIGVFTGMNSQYSDCDPLGNTYAPIYL-----RKPQD 561
 Db 2738 ---IRLTOQKGEATHN--GV-----QIANDAPFOLISVEDPELFVQRELAP 2778
 Qy 562 QTEAAKATMDTYRAT-LEPLF-----IDLEOPRLLDGPASS--BGL 602
 Db 2779 QLCQSFETILDVYPATOMQRFELINPVTKPRSRSTPFEHDPPD-----ADCAASLMRAC 2832
 Qy 603 SSVIVDHPTRRILDTLRLRIEQFTTQEMKVLUVTRDKTIREGSEATHS----- 652
 Db 2833 ASLAKHFDITFVFLEARGLYQVVLKHVDPIEM--LOTENINSASTRSPFLDVDAEKPI 2890
 Qy 653 -----MALTDFDYSGAFCPITNFYKRTTHAVQDIALS---QCHCFYQGQQV-BG 699
 Db 2891 RLGQPLLITALLERPGS-----TLRVLRILSHLILHSLHILFGGSLLPP 2944
 Qy 700 RNFERNQFOQVLU--RRRFYDYLENGGPISTRSITITLSECPVSAPNPTQGQDAPAGRFTDGD 757
 Db 2945 RPKAGYMQHASSREGDFFWR-SVLRDSSMTYIKGNNNNTTPPPQQQSTS5GGAHASK 3003
 Qy 758 LARYSVEVTRDVKRNVVPSGNCT-----ILSEARRARLY----GLASAYIQ 800
 Db 3004 VVTLPTQANTDSR-TRATFTTACALMLAKEIDNSDVVFGRTVSGRQGLPLAHQ 3057

RESULT 13
 AAY00206
 ID AAY00206 standard; Protein; 1074 AA.
 XX
 AC AAY00206;
 DT 20-APR-1999 (first entry)
 XX
 DE Enterococcus faecalis protein BP102.
 XX
 KW Enterococcus faecalis; infection; vaccine; immune response; diagnosis;
 KW detection; attenuation; antigenic.
 XX
 OS Enterococcus faecalis.
 XX
 PN WO9850554-A2.
 XX
 PD 12-NOV-1998.
 XX
 PP 04-MAY-1998; 98WO-US08959.
 XX
 PR 14-NOV-1997; 97US-0066009.
 PR 06-MAY-1997; 97US-0044031.
 PR 16-MAY-1997; 97US-0046655.
 XX
 PA (HONDA) HUMAN GENOME SCI INC.
 XX
 PI Bailey C, Choi GH, Hromockyj A, Kunisch CA;
 XX
 DR WPI: 1999-070095/06.
 DR N-PSDB; AAX20196.

PT New isolated Enterococcus faecalis polynucleotides - used to develop
 PT products for the detection of Enterococcus and for use in vaccines
 PT for prevention or attenuation of Enterococcus infection
 XX
 Claim 9; Page 197-198; 301pp; English.
 XX
 CC The present sequence represents a protein isolated from
 CC Enterococcus faecalis. The present invention describes genes, proteins
 CC and antigenic polypeptides isolated from E. faecalis. The proteins can

be used in vaccines for preventing or attenuating an infection caused by a member of the Enterococcus genus in an animal. They can also be used for detecting Enterococcus antibodies in a sample. The nucleotide sequences can be used for detecting Enterococcus nucleic acids. Products from the present invention can also be used for screening compounds to identify agonists and antagonists of E. faecalis protein activity.

xx Sequence 1074 AA;

Query Match Score 123; DB 20; Length 1074;

Best Local Similarity 18.3%; Pred. No. 0.11; Matches 206; Conservative 150; Mismatches 415; Indels 352; Gaps 52;

Db 250 RVNNTAVARGAAHLAFAFDENHEGAVLPPDITYTFOSSSSGTTARGARRDNVTSKPS 309

67 RTTSIYAETNGAKQTVFV-IEPVGSIPIPEVTHGY-----QRNPL 104

Qy 310 PGGFERRLASIM-AADTAHLAEVINTGIYPE-----TPTDIKERPMPI 353

Db 105 PSMSDKAKLVSLWMEKAGTDIDINMVAQKMIWEDVNGYKLHSIKRLLGASYDVK-----QRNPL 104

Qy 354 GMECTLPRNALGSYTYAR-----VAGVIG-AMVFSPNALSYLTVEVDGSMTEAKDGGP 405

Db 159 SIEBKINK-AIEEYQKCFSPFHNTTVKTHLQGTTLDKNEELNSERDKVQNTA-----211

Qy 406 GPSPNRFYOFAGPHLAANPQTDRGHGVLISSQSTSNTESVYDYLALICGEGAPLIALRL 465

Db 212 ---NIDRVIGNQLVLTP-----NSNSKSGTLILKKSAGTGTP-VAYKK 251

Qy 466 FYLERCDAGAFTGHDGDLALKYTGTGDFSBIPCSCLCEKTRPVCAHTTVHRLRQRMFRFGQ 525

Db 252 AGLOQTMAGALDKEPTNTYAIKINNTETKGS-LKIKDIKESEGDIVPETVPHL-----DFGK 304

Qy 526 A-----TROPICUVFGTMNSQYSDCPDGNTA-----PYLILRKPGDQTEAKATM 570

Db 305 ALPSKDVTIDKGQI-----SILDGHGHTKVTITEKSVFDPYMDDTPPAATITAGETI 358

Qy 571 QDITYRATERLFLIDLEQERLILDRGAPSCSEGLS-----SYTVDHPT--FRRILDILRA 621

Db 359 SMPSKRNDRQKGQILLERG-VETGTDLWNRDYSLAANTFAIRDSEPAGEVQETTDEKG 417

Qy 622 RIOTTOTFQMKVLUETRDYKIREGLSEATHSMALTDP-----YSGAFPITFLVURT 675

Db 418 RAE-TPELANALELSTYYTE--TKSSNGFVFTKPTKVELKYANOTVALTSNVXKGQ 473

Qy 676 HUAWVQDIAL-----SOCRCHFYQGQV-----EGRNFRNQFPVLLRRFVDL 717

Db 474 NQITGTGTLTKEDDTGNEQSQQKAEGSYTFTAKDGOAVKNSERAK-----TEL 526

Qy 718 FNGGFISTRISITVILSE-GPVNSAPNPTLGQ-----DAPAGRFTGDLARVSVE-----764

Db 527 VRKTAKADETTVLADEKNOVAKKHIAINEWQETKAPAGTYLDETRKPVSIKVUDNE 586

Qy 765 -----VIRDYRNRRV-----FSGNCNTNLSEAR-----PARLYGLASAYOROKR 805

Db 587 KNAVIREDTVTAKEQVTRFGDFEFFKFGSADGTAEGNDLUSFKSPLPLECTXEITGAEDIA 646

Qy 806 VDMHLHGALGF-----LIKQFHGLLPRGM-----PNSKS-----E-NP 838

Db 647 TTACNEBQJLGFQGYKFENLPYDYLLEIEA---BEGFQKITPLFIRSTKENDDYAKS 703

Qy 839 QWFWTLLQRNO-----MPADKLTHEBITIAAVKRFEEYAINFLP-PTCIGE 888

Db 704 EYFTTIEEGQKPIKNTVYECUTTNNE-----FSVSLNRLMLYDPEKEDSITS 754

Qy 889 LAQFYMANLILKCYDHSQYQINTLTSITGARRPDPSVLMWIRKD---VTSADIBTQ 945

Db 755 LATWDGCKKCNLDFTE-LVDKL-----RYNLHEKDVYVAQDIVEA- 799

Qy 946 AKALLKENTNLPELWTAFTST-----HLVRAAMNQRMVYLGISISKYHGA 992

Db 800 TKAQSKDEKAKPVVIAETATLANKETGTWKHLHKLTAEQ-----VLDKSVLFNYV 853

CC be used in vaccines for preventing or attenuating an infection caused by a member of the Enterococcus genus in an animal. They can also be used for detecting Enterococcus antibodies in a sample. The nucleotide sequences can be used for detecting Enterococcus nucleic acids. Products from the present invention can also be used for screening compounds to identify agonists and antagonists of E. faecalis protein activity.	Qy 993 AGNNRYFOQANWSGLNG3KNUCPLFTEDRTRRFIACPRGGFICPYTGSPSGNRETTLSD 1052
xx	Db 854 YENKAYFAEIGNE-----PVA-----KDA-----
xx Sequence 1074 AA;	Qy 1053 QVRGIIVGSGAMVOLAIYATVRAVGAR-----AQHMADFDDWLSITDDEFIARDL-----1102
Query Match Score 123; DB 20; Length 1074;	Db 876 QAQ-----TVNCTIERHVSIGTQTAHLEDGSOTFTHGDMMDFDVSVTHDVLD 923
Best Local Similarity 18.3%; Pred. No. 0.11; Matches 206; Conservative 150; Mismatches 415; Indels 352; Gaps 52;	Qy 1103 --EHDHDQITQLETPWVYGALEBAVKE-----KTAGD---GETPTLNANFD 1149
Db 67 RTTSIYAETNGAKQTVFV-IEPVGSIPIPEVTHGY-----QRNPL 104	Db 924 GSKEAPETIYALLPDGTMREBIWYSGKLEHNDKEFIFTKTAEVKDTGKYPEGTKEF- 982
Qy 310 PGGFERRLASIM-AADTAHLAEVINTGIYPE-----TPTDIKERPMPI 353	Qy 1150 SCEPHTDTISNV-----LNISSGNISGSTVPLKRPEDDE 1185
Db 105 PSMSDKAKLVSLWMEKAGTDIDINMVAQKMIWEDVNGYKLHSIKRLLGASYDVK-----QRNPL 104	Db 983 -TEINYEKDGVNVNGKHNEDLKERQSTLTPKEVPTIPISTPKQQPS 1024
RESULt 14	
Qy AAY00188	Qy AAY00188 standard; Protein; 1074 AA.
Db XX	Db XX
Qy AC AAY00188	Db DT 20-APR-1999 (First entry)
Db XX	Enterococcus faecalis protein EF094.
Qy XX	Enterococcus faecalis; infection; vaccine; immune response; diagnosis; detection; attenuation; antigenic.
Db XX	Enterococcus faecalis.
Qy XX	WO850554-A2.
Db XX	PD 12-NOV-1998.
Qy XX	PF 04-MAY-1998; 98WO-US08959.
Db XX	PR 14-NOV-1997; 97US-0066009.
Qy XX	PR 06-MAY-1997; 97US-0044031.
Db XX	PR 16-MAY-1997; 97US-0046655.
Qy XX	(HUMA-) HUMAN GENOME SCI INC.
Db XX	PJ Bailey C, Choi GH, Hromockyj A, Kunisch CA;
Qy XX	WPI; 1999-070095/06.
Db DR N-PSDB; ARX20178.	
Qy XX	New isolated Enterococcus faecalis polynucleotides - used to develop products for the detection of Enterococcus and for use in vaccines for prevention or attenuation of Enterococcus infection
Db PT	PT Claim 9; Page 186-187; 301pp; English.
Qy XX	The present sequence represents a protein isolated from Enterococcus faecalis. The present invention describes genes, proteins and antigenic polypeptides isolated from E. faecalis. The proteins can be used in vaccines for preventing or attenuating an infection caused by a member of the Enterococcus genus in an animal. They can also be used for detecting Enterococcus antibodies in a sample. The nucleotide sequences can be used for detecting Enterococcus nucleic acids. Products from the present invention can also be used for screening compounds to identify agonists and antagonists of E. faecalis protein activity.
Db CC	CC Sequence 1074 AA;
Qy CC	CC Query Match 2.0%; Score 123; DB 20; Length 1074;
Db CC	CC Best Local Similarity 18.3%; Pred. No. 0.11; Matches 150; Mismatches 415; Indels 352; Gaps 52;

250	RVRNTAVARGAAHLADEFNHEGAVLPPDITYTQFQSSSGTTTARGARRNDVNSTSKS	309	Db	983	-TEINYERDGNYNGKHNEDLKEKSQTLTIPKVEPTIPSTPKOPE	1024
67	RTTSIYAENGAKQTVC-IEPGVSIPTFVHG-	104				
310	PSGGFERRLASIM--AADTALHAEVTFNGYFE-----TPDIKEWMPF1	353				
105	PSMSDRKLVSIVMEKAGTDITNMVAQRMWVHKGSSVTK-----	158	RESULT 15			
354	GMEGTIPRNALGSYTA-----VAGVIG-AMVFSNSPAVLVLTEDSGNTAEAKDGGP	405	ABP43407			
159	SIEGKINK-AIEEYQKPKPSFHNTTVKTILGOSTTLDKNENLNSEPKVUQNTA-----	211	ID ABP43407 standard; Protein; 1074 AA.			
406	GPSNRNFQFAGPHLAANPQTDGHLSSQSTGSSNTFSDYLALICGPGAPLLARLL	465	XX			
212	--NIDRVIQNQLVLTP-----NSNSKGSLTILKKSAGTGT-VAYKK	251	AC ABP43407;			
466	FYLERCDAGAFTGHDALKYVTGTFDSEIPLCSLCBGTTRPVCAHTVHLRQMRPFGQ	525	XX			
252	AGIQTVMAGALDKPNTAIKINVETKGS-LKZKKIDKESGDDVPTVFHL-----DFGK	304	DT 05-AUG-2002 (First entry)			
526	A-----TQPQIGVFGTMNSQYSDCDPLGNYA-----FYLIILRKPGDQEAAKATM	570	XX			
305	ALPSKDVTDTDRG-----SFLDGPBPGTKVITKTSVPDYMIDTPTMMATIKGETI	358	XX			
571	QDTYRATERLRLDIFLDEQRLLRGAPCSSEGUS-----SVTVDHPT---FRILDTLRA	621	XX			
359	SNTSKNNRQKGQILLEKTG-VETGTDLWNDDNSLAGNTFAIRKDPSAGEIVETITDEKG	417	XX			
622	R*EQTTQFMKVLVETRDYKAREGLSBAATHSNALTFDP-----YSGAFCPITNFLVKRT	675	XX			
418	R2E-TKEELANALELGTYTYE-TKSSNGFVNTPKPTKVELKYANQTVLVTNSVKQG	473	PI Choi GH, Bailey C, Hromockyj A, Kunsch CA;			
676	HIAVQQDIAL-----SOCCHCYFGOOY-----EGRNFPNOQFQVFLRRRFVDL	717	XX			
474	NQBITGBTTLXEDKTGTGNESOGKAEFKGAEVTLFTARDGOAVKNSEAFK-----TEL	526	XX			
718	FNGGFISTRSTVYUSe-GPY/SAPNPFLGQ-----DAPAGRTEGCDLARVSE-----	764	XX			
527	VKGTKRASDETVLALDEKNOVQAHKHLAINEFWQOETKAPECTYLTDETGYKPVSIKVDNN	586	PS Claim 9, Page 163; 255pp; English.			
765	----VARDIRVKGRTV-----FSGNCINLSEAA-----RARLVGLASAYQRQBR	805	XX			
587	KNAVIRDVTDAKEQVTRFGDFEFFKFKGSAODTAEFGENDLSFKVSPLEXTKEIGAEDYA	646	CC The present invention provides the protein and coding sequ-			
806	VDMHLHALGF-----LRLQFHGLFPRGM-----PPNSKS-----PNP	838	CC number of polypeptides from Enterococcus faecalis, use-			
647	TTACNQLQGFQGYKGENLPYGDYLLEIEA---PEGFQKITPLFIRSTFKENKDDYAKS	703	CC pt vaccines for preventing or attenuating an infection			
839	QWFWTLLQRNO-----MPADKCLTHEEITTAIAVKRFEEYAAINFNL-----PTCIGE	888	CC by a member of the Enterococcus genus in an animal, particula-			
704	EYVFITEEGOKQPKMVTVFYEKUTNNE-----FVSLSLRMLYDLPKEDSDTS	754.	CC rly faecal infection. The protein is also useful for preventing or treat-			
889	LAQFYMANLILKYYCDHSQYLINTLTSITGARRPRDPSSTLHWIRKD--VTSADIEIQ	945	CC ing the present sequence is a protein of the invent-			
755	LATWDGNKLINTLDE-TVD KL-----RYNLHELKEDWTVVQAIDVEA-	799	SQ Sequence 1074 AA,			
946	AKALLEKTENLPEMNTAFTST-----HLVRAAMNORPMVTLGIISTSKYHGA	932	Query Match Score 123; DB 23; Length 10;			
800	TKAQAEKDEKAKPVVIAETTALANKEKIGTWKLHKLTAEQ-----VLDKSVLFLNV	853	Best Local Similarity 18.3%; Pred. No. 0.11;			
993	AGNRFVQAFNWSGLNGKNVCPLFTDRTRPFIACPRGGFICPVTGESSGRRETTSD	1052	Matches 206; Conservative 150; Mismatches 415; Indels 1;			
854	YENKWAERGNE-----VNGTIEHRSISQKAHLEDGSQFDTDEFLARDL-----	923	OY 250 RYRNVTAVARGAAHLADEFNHEGAVLPPDITYTFOSSSSOTTGARTGARI			
1053	QVRGIIVGSSGAMVQALIAYATVVRAYGAR-----AOHMAFDWLSLTDEFLARDL	1102	Db 67 GMGETGLPRLNALGSTAR-----VAGVIG-AMVFSNPNALYLTBEVED			
876	QAO-----TVCNTERHRSISQKAHLEDGSQFDTDEFLARDL-----	923	Db 159 SIEGKINK-AIEBYQKPKPSHNTYVKTIGQSTTLIDKREBLNLSSEFDKV			
1103	--EELHDOLIOTLETPTWVTEGALAVKILDE-----KTTAGD---GETPMLNATNF	1149	QY 310 PSSGFERRLASM-M-AADTAJHAEVTFNTGTYE-----T			
924	GSKRAFETTLYALLPDGTNKEWWSKGKSHHEVNDKEFTVLAKEVDTGKYPGTTKTF-	982	Db 105 PMDSKAKLVSFLWMBRGTDDTNMVAQKTIWEYNGKHLISKRIGGAA			
1150	SCEDSHDTSNSV	1150	QY 354 GMEGTGLPRLNALGSTAR-----VAGVIG-AMVFSNPNALYLTBEVED			
			Db 212 FYLERCDAAGATGCGGDALKYVUTGTFDSEIPCSLOCEKHTPVCANTHTVR			
			Db 466 252 AGLOQTMAGALDKPNYIAKINVERKGS-LIKKIKLDEKODDIVPETVFL			

Qy 526 A-----TROPICVGFTMNSSOYSDCDPLGNYA-----PYLILRKPGDOTEAAKATM 570
 Db 305 ALPSKDVTDDKG-----SILDG3PHGKXVITTEKSVPDPMDTPMAATIKAGETI 358
 Qy 571 QDTYRATLERLFDIEBQLDRGAPCSSEGLS-----SVYDHEPT---FRIILDTIR 621
 Db 359 SMTSKNMRRQGQILLEKTG-VETGTDLWNDNYSLAGNTPAIRDSPAGEIVQDITDEKG 417
 Qy 622 RIECQTTOQFMVKLVETRDYKIREGLSEATHSMALTEDP-----YSGAFCPITNEFLYKRT 675
 Db 418 RAE-TPKELANALELGTYRTB-TKSSGFNTFKPKVLEKXANGTVALYTISNTNRKGQ 473
 Qy 676 HLAIVQDIAL-----SQCHCFYQGQV-----EGRNFRNOQFQPVLRFFVDL 717
 Db 474 NQEIGETTLTKEKDGTGNESQGAEFKGAEYLTFTAKDQAVTKWSEAFK-----TEL 526
 Qy 718 FNGGELSTRSTTIVLSE-GPVSAANPTLCQ----DAPAGRPFDDGALARVSVE-----764
 Db 527 VRGTRASDEVTLALDERKQVAQHLLAINEFWQETRAPEGYLTDETKYPISKYDNNE 586
 Qy 765 ---VIRDIVRKNRUV-----PSGNCTNLSEAA-----RARLYGLASAYQROEKR 805
 Db 587 KNAVITRDYTAKEQYVIRFSEFFAGSADGTNETGENDLSPVSPLEGETXBTGAEKA 646
 Qy 806 VDMUHGALGE-----LLKQPHGLLFPRGM-----PNPSKS-----PNP 838
 Db 647 TTAINEQOLGFDGYKGFENLUPGYDYLLESEA---PEGFQKITLETRSTFKENDDYAKS 703
 Qy 839 QWFVILLQNO-----MPADKLTHEBITTAIAVRKFTEYYAAINFNP---PRCIGE 888
 Db 704 EYVTTITEGQKPIKMTVPELTINNE-----FSVLNRMLYDLPKEBEDSLTS 754
 Qy 889 LAQFYMANLILKYCDSQYLNINTLSITIGARRPRDSSVLUHMIRKD---VTSAADETQ 945
 Db 755 LATWDGDKNLNLDFTL-LVDDI-----RYNHLIKEKDWYVVAQALDVEA- 799
 Qy 946 AKALLEKTNLPELWTTAFTST-----HLYVRAMNNQRPMVVLGJISISKYHGA 992
 Db 800 TKAQEQDEKAKVVIATTANKEKTMKILKLTAEQ-----VLDKSVLFLNYV 853
 Qy 993 AGNMRVFOAGNWQGLNGKNCVCLFTDRTRFLIAPRGGFTCPYGSSGNRFTLSD 1052
 Db 854 YENKVAFFAGNE-----PVA-----RDASLNN 875
 Qy 1053 QVNGIISSGAMVQLAYATVRAVGR-----AQHMAFDWLSSLTODEFLARDL-- 1102
 Db 876 QAQ-----TINCTIHPHSVOTKNALEDSQTFHGDMMEDDVSVTHDVLD 923
 Qy 1103 --BELHDQIQTLETPTWVEGALEVILDE-----KITAGD---GETPTNLAENFD 1149
 Db 924 GSXEAFFTILYALLPDGTNKEIWGSKIEREVNDKEFTKTVLAKYDTCRYPEGTFKFTP- 982
 2y 1150 SCPSPHDTSNSY-----LMSGSNTSGSTYPLKRPPEDE 1185
 Db 983 -TEINYERDGNYKGHNEDLKERSQUTLPKEVPTIPSTPKOPE 1024

Search completed: January 30, 2004, 13:12:34
 Job time : 52 secs

GenCore Version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

protein - protein search, using sw model

on: January 30, 2004, 13:14:08 ; Search time 44 Seconds
 (without alignments)

file: US-09-679-609-2
 Direct score: 6394
 Sequence: 1 MENTQKTVTVPGLGVVAA..... DELFDLSGIPKHNITEM 1203
 Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 79580 seqs. 207824079 residues

Total number of hits satisfying chosen parameters: 789580

hit-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

abase : Published Applications_AA:
 1: /cgn2_6_ptodata/2/pupbaa/US07_PUBCOMB.pep.*
 2: /cgn2_6_ptodata/2/pupbaa/US07_NEW_PUB.pep.*
 3: /cgn2_6_ptodata/2/pupbaa/US06_NEW_PUB.pep.*
 4: /cgn2_6_ptodata/2/pupbaa/US06_PUBCOMB.pep.*
 5: /cgn2_6_ptodata/2/pupbaa/US07_NEW_PUB.pep.*
 6: /cgn2_6_ptodata/2/pupbaa/US07_PUBCOMB.pep.*
 7: /cgn2_6_ptodata/2/pupbaa/US07_PUBCOMB.pep.*
 8: /cgn2_6_ptodata/2/pupbaa/US08_PUBCOMB.pep.*
 9: /cgn2_6_ptodata/2/pupbaa/US09_PUBCOMB.pep.*
 10: /cgn2_6_ptodata/2/pupbaa/US09B_PUBCOMB.pep.*
 11: /cgn2_6_ptodata/2/pupbaa/US09C_PUBCOMB.pep.*
 12: /cgn2_6_ptodata/2/pupbaa/US09_NEW_PUB.pep.*
 13: /cgn2_6_ptodata/2/pupbaa/US10_PUBCOMB.pep.*
 14: /cgn2_6_ptodata/2/pupbaa/US10B_PUBCOMB.pep.*
 15: /cgn2_6_ptodata/2/pupbaa/US10C_PUBCOMB.pep.*
 16: /cgn2_6_ptodata/2/pupbaa/US10D_PUBCOMB.pep.*
 17: /cgn2_6_ptodata/2/pupbaa/US60_NEW_PUB.pep.*
 18: /cgn2_6_ptodata/2/pupbaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	3054	48.5	1452	14 US-10-050-673-2	Sequence 2, Appl1
2	3015.5	47.9	1196	12 US-10-050-562-200	Sequence 200, App
3	3015.5	47.9	1196	12 US-10-037-551-200	Sequence 200, App
4	3015.5	47.9	1196	12 US-10-237-551-232	Sequence 232, App
5	788	12.5	274	12 US-10-337-551-231	Sequence 231, App
6	736	11.7	248	12 US-10-337-551-230	Sequence 230, App
7	123.5	2.0	850	15 US-10-156-781-9121	Sequence 912, App
8	120.5	1.9	2597	10 US-09-905-129-12	Sequence 2, Appl1
9	120.5	1.9	2597	10 US-09-905-129-10	Sequence 10, Appl1
10	120.5	1.9	2597	10 US-09-905-129-13	Sequence 13, Appl1
11	120.5	1.9	2597	10 US-09-991-630-2	Sequence 2, Appl1
12	120.5	1.9	2597	10 US-09-991-630-10	Sequence 10, Appl1
13	120.5	1.9	2597	10 US-09-991-630-13	Sequence 13, Appl1
14	120	1.9	2597	12 US-10-084-146A-8	Sequence 11, Appl1
15	118	1.9	19608	12 US-10-084-146A-8	Sequence 8, Appl1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

סודות נסיך

TYPE

ORGANISM: herpesvirus

US-10-050-673-2

卷之三

BOSTONIAN

Matches 609: Cancer

卷之三

MENTOKTVT

METTURKUTTAI

卷之三

56 KTFESSLAV

卷之三

39 SGF EANVAV

0v 111 6 ABE BEG FSB

卷之三

119 ARRHFGEFD

卷之三

卷之三

卷之三

Qy 176 GGLDVTHTNHDVTRIPLFPVQLFMPDVYRNLVPDPENTHRSIGEGTVYPFPYNTGLCH 235
 Db 179 GSDXKVTGGAETHRIPYPLQMPDSRVLFPVANHRSGTENITYPFPFNRLPQR 238
 Qy 236 LHMDCVIAPMAVALVRVNTAVARGAHLAPDNHEGAVLPPDITYFOSSSGTTAR 295
 Db 239 LFFIAVVAALVRCRNTDAVRAAAAHAFDNHEGAALPADITTAFAEASQG-KTPR 296
 Qy 296 GARDNDVTSKSPSGGERIASIWADTAHAETFTNTYEPPTDKWPNFMIG 355
 Db 297 GGR----DGGGKPGAGSFEORIASWMDAALLESIVSHAVDFEPTDSAWPLCEQQ 351
 Qy 356 EGJGPRLNALGSTARYAVGIVGMFSSENSALYLTEVDGTMTEAKDGPROPSFRYQF 415
 Db 352 DTHAABANAVGALARAGLVGNVFTINSALHEDVDAAPDHSK-PSFTRFLV 410
 Qy 416 AGPHLAANPQTDGDHVL-----SSOSTGSNTFSVDYLALIGFGAPLLAULPVL 468
 Db 411 PGTHVAAPQVDIEGHVPGFBORPTPLVGCTQ-EPAHENLAMLGCFSPALLAULPVL 469
 Qy 469 ERDAGATGGCG-DALKYVTTGSDSSAIPCSCLCEKTRPVCAHTTHRFORMPFGQT 527
 Db 470 ERDGGVTVGRQMDVRYVADSNOTUPCNLCTFDTRHACVHTTMLPARHPFASAA 529
 Qy 528 RQFIGVFTMNSQYSDODPLNYAPYLRLRGDQEBAKATMDTYRATLERLTDLQ 587
 Db 530 RGRIGVGTNEMYSDCDVLGNYAAFSALKP-ADGETARTIMDLYRATERMNEELT 588
 Qy 588 ERILDGRAPCSSEBGLSSVIVDHPPTFRILDTURARBOTTOFQKVLVTRDVKIREGLS 647
 Db 589 LQYDQAVPTANGRLLEITITREALHTVNNNVRQYDREBQELMNLVEGRNFKFRDGLG 648
 Qy 648 EATHSMALTFDYSIGACPITNLFVLRDIALSOCHCYYGOOYEGRFRNFOQ 707
 Db 649 EAHAMSLTLDYACGPCLQLQGRSRNLAVYDALSQCHYFAGQSTEGRFRNFOQ 708
 Qy 708 PYLRRRFVDLFNGFISTRSTVUSEG-PVSAPMPTLGDAPGRTEGDLARVSVEV 766
 Db 709 PYLRRRYMDMFNGFUSAKTIVAUSEGAAICAPSITAGTAPESSEFQDGVATLGFPP 768
 Qy 767 RDIRVKNRVVFSGNCNTNLSEARALVGLASAXQBEKRDMMGALGFLKOFGLJPP 826
 Db 769 KLRVSRVLPAGASANASEAKARYASQAYPKDKEVLDLGPLGLKOFHAAIFP 828
 Qy 827 RGMPPSKSPNPQWMTLQONQMPADKLTHEETIAAVKRPEYAAINFNLPPICTI 886
 Db 829 NGKPPESNOPIQWFTALQRNQPLRSDEETIATPKISDGAINEFLNLPPICTI 888
 Qy 887 GELAQYMANLILKYCDHSQWLINTLTSITGARRPRDPSSTVIRKDVTSAADIEQ 946
 Db 889 SELAMYMYMAQILRYCDHSTYFINTLTAGSRPSPSYQAAAAM--SAQGGAGLEZKA 945
 Qy 945 KALLEXTENLPFLWTAFTSTHLMAMNORMPMVYGLTISKYGAAGNNRVEQAGNWSG 1006
 Db 946 RALMDAVIDHAWGATSMFASCNLLEPMAMARPMVGLGUSISKYGMANDRVEQAGNWS 1005
 Qy 1007 LNGKRNVCPLPFTEDPTTRFETIACRGGECPVTGPGSSCNETTISDQRTGIVUSGGANQ 1066
 Db 1006 LMGKNAACPPLIIFDTRKFVACRPGFVAASTLGGAAHESSICEORGISEQRTGIVUSGGANQ 1065
 Qy 1067 LAIATVVRAVGARAHQMAFDWLSLTDEFLADLDEEHDQITQLETPWIVEGAL--- 1123
 Db 1066 SSVFVATVKSGLPRTQQLQEDWALLEBEYLSEMMELTARAERGENGWTDAALEVA 1125
 Qy 1124 -EAVKILDEKTTAGDGETPNLNANFD--SCEPHDITSNVNTISGMSISGTPGJKRP 1180
 Db 1126 HEAEALVSOQNGAEE----VNFNGDFGCE---DDNATPFGGCAAPAGRKRA 1173
 Qy 1181 PEDDELFDSLGSIPRKHGNITMEM 1203
 Db 1174 FHGDDPPG-BGPPDKKGDIILDM 1195
 Qy 1174 SLSIDPYTCGPCLPOLLARRSNLAVYDLSQCHVTPYGOOVERNFRNQFQPVJRR 713
 Db 655

RESULT 2
 US-10-200-562-200
 Sequence 200, Application US-10200562
 Publication No. US20030165819A1
 GENERAL INFORMATION:
 APPLICANT: McGowen, Patrick
 APPLICANT: Hosken, Nancy A.
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
 FILE REFERENCE: 10121-538C2
 CURRENT APPLICATION NUMBER: US/10/200-562
 CURRENT FILING DATE: 2002-07-19
 NUMBER OF SEQ ID NOS: 212
 SOFTWARE: FastEQ for Windows Version 4.0
 SEQ ID NO: 200
 LENGTH: 1196
 Score 3015.5; DB 12; Length 1196;
 Best Local Similarity 49.0%; Pred No. 4e-290;
 Matches 595; Conservative 213; Mismatches 366; Indels 41; Gaps 16;
 Query Match 47.9%;
 Score 3015.5; DB 12; Length 1196;
 Best Local Similarity 49.0%; Pred No. 4e-290;
 Matches 595; Conservative 213; Mismatches 366; Indels 41; Gaps 16;
 Qy 4 TQCTVYETGPYGYVY-ACKVEDDILEEFLARSTSDSLALLPMLNLTVEKTFTSS 61
 Db 7 TTTTVKTPPGPGYVGRACAEGLL-LLLSARSQGDAVAVAPLIVGLTVESGFEAN 64
 Qy 62 LAVVSGARTTGHAGAGITLKUTTSHFYPSPVPRGKENTPSSAAPPNLTACNARBERG 121
 Db 65 V3AVVGRGTRTGTTGATVSLKIMPSPYSPVPHGRHIASTSTOPNLTLCERAPRPHG 124
 Qy 122 FRCQGPFDVQETGATEOTRIGLEPNTILKIVTIALEKAEVNCFNVLHGGLDIV 181
 Db 125 FADYAPPDCDQKHEETGDAICERLGIIDPDRALYLVITSGREAVCISNTFLHGGMDKV 184
 Qy 182 HINGDVIRIDOLPFVPLMPVYRNAYPDPEPNTHERSISGEGFVYPTPFTNTGLCHLHDGV 241
 Db 185 FTGDAYPHRIVTYPLQEMDPSRVIAFDPMCNHRSIGENFNFLPFFERPLRLLFEAV 244
 Qy 242 ZAPMAYALVRYNVNTAYARGAHLADEFNTEGGAVIPPDITYTFSQSSSGTTARGARRND 301
 Db 245 VCPAAVALRANNUDVARAHHLADENBEGAAIPADIETFAFEASQG-KPQGAR--- 299
 Qy 302 VNSTSKEPSGFGFERIASIMAATAUHLEVWENTGILIEETPTDIKEWMEIGNEGLPR 361
 Db 300 -DAGNKGPGGFEORLASMAGDAALAESIVSMAVEPDPPPPITTPPLLEGOTPAR 357
 Qy 362 INFALGSYTARYAGVIGAMYEPSNSALTYTEVEDSGMTBKDGGRGPPSNRFQOPAGHFLA 421
 Db 358 AGAVGAYLARAAGVIGAMYESTNSALHTEVDDAGPDKDHEK-PSYVRFELVPGHVA 416
 Qy 422 ANPQTDRGHVL-----SSQSTGSSNTFSDYVLAJICGFSAPLARLITYLERCDAQ 474
 Db 417 ANPOLDREGHVPGYEGRTAPLYGGTO-BRGEHHLAMCGSPALLAKMFLYLERCDG 475
 Qy 475 AFTGCHG-DALKYVGTGTRSEIPPSLCLRKRTPVCAHTVHLRORURPFGCATROPIGV 533
 Db 476 VIVGRQEMDVRYTAIDGQDVIDVONLCFTETRACHTTLMFRARPKFASAARGAIGV 535
 Qy 534 FTGTMSONYSCDPIGNAYYLTKPGDQTEAKATQDVTATLELFIDIEQERULLDR 593
 Db 536 FTGTMSONYSCDVGANYAAFSALKR-AUGSFRTINQSTYRATERVMATEALQVVDQ 594
 Qy 594 GAPCSSEGIVSIVDHPPTERRLDTLRARIEOTTQPKVULVETDVKIRCLSEATHSM 653
 Db 595 AVPTALGRLETIGNREALHTWNNNIKOLVDESEQMLNLBGRNFKFRDGLAEANHAM 654
 Qy 654 ALTPDPSGAFCPITNFVLYKTRILAVVQDLSQCHVTPYGOOVERNFRNQFQPVJRR 713
 Db 655 SLSIDPYTCGPCLPOLLARRSNLAVYDLSQCHVTPYGOOVERNFRNQFQPVJRR 714

714 FVDLFNGGGFISTRITYTLESEG-PVSAPNPTLGGDAPAGRTEPDSDLARYSVEYTRDIRVK 772
 715 VNDLFNGGFLSAKLTVALSEGAIACAPSLTAGGTAPEASSFEEDVARTLGPKELRVK 774
 773 NRVVFSGNTNLSEARLVGLASAYQOEKRVDMLRHALGFILKOPHGLLPRGMFPN 832
 775 SRVLFGASANASEAKARASQYQPKDVRDILGPGLFILKQTHAVIPNGKPPG 834
 833 SKSPNPQWFTLLOQNQMPADKLTHEITIAAYTRFBEYAAINFNLPLPTCIGELAQF 892
 835 SNQNPPOWFETALQRNQLPARISSREDIETIAFKRSFDYGAINFNLAPNNSELAMY 894
 893 YMANLILKYCDHSQYLINTLTSITIGARRERPDPSVLEWIRKDVTSAADIEOQAALLEK 952
 895 YMANSQILRYCDHSTTFINLTAVAGSRREPSVAAAAWAPQ--GGAGLEAGARALMDS 951
 953 TENPELWTTAFTSTHLVRAAMNORPMVYLGISISKYIGAAGTNTRVEAGNWGLNGKN 1012
 952 LDAHNGAWTMFASNLLEPVMARPMVYGLSISKYGMAGNDRVFOAGNWASLUGKN 1011
 1013 VCPLFETDRTRRFETACPRGGFICPVTPGPSSGNETTISDQVRGIIYSGGANVQALIYAT 1072
 1012 ACPLLIFDTRKFVLACPAGFVGAASSLGGGHEHSLQEQLRGIIAEGLAAVASSFVA 1071
 1073 VRVARGAROHMADDWLSTDDBLARDLEELDQIOTLTLPTWVEGAL--EAVKI 1128
 1072 TVKSISQGPRIOQLOEDWNLALLEBLYSEMMETPTRALEGHEWSDAALEEVAREAL 1131
 Y 1129 LDEKITAGDGETPTLNANFDSCPSHDITTSVNLNISNSNISSTVPCLKRPEPDBELFD 1183
 b 1132 VSQGLAGE----VFNFGDFGEDDHAAFSFGGLAA--AGAAGVARKRAFHGDPPG 1182
 Y 1189 LSGPIKHGNITMEM 1203
 b 1183 -EGGPERK-DLTLDM 1195

RESULT 3
 S-10-237-551-200
 Sequence 200, Application US/10237551
 Publication No. US20030165820A1
 GENERAL INFORMATION:
 APPLICANT: Day, Craig H.
 APPLICANT: Hosken, Nancy A.
 APPLICANT: Parsons, Joseph M.
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
 FILE REFERENCE: 210121:58C3
 CURRENT FILING DATE: 2003-09-06
 NUMBER OF SEQ ID NOS: 254
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 200
 LENGTH: 1196
 TYPE: PRT
 ORGANISM: HSV2
 S-10-237-551-200

Query Match 47.9% Score 3015.5; DB 12; Length 1196;
 Best Local Similarity 49.0%; Pred. No. 4e-290;
 Matches 595; Conservative 213; Mismatches 366; Indels 41; Gaps 16;

Y 4 TOKTUTVPRGPGKYY-ACRVEDLDLEBISLALARSTDSDIALPLMRNLTVEKPTFSS 61
 b 7 TTRVKVPQGPQMVYGRCPAEGLEL-LSLISARGDADYAVAPLIVGTVVESEPRAN 64

Y 62 LATVSGARTGLAGAGITKLTSHTFSPVYERGGKHVLPSAAPNLTRACNAARERFG 121
 b 65 VRAVGSRTGLGCGTAVSLKLMPSHYSPSVYFHGRHLASTQANLTRICERPHFG 124

Y 122 FSRQGPPVYDGAVETTGABEICTRGLEPENTILYVTAFLKEAVAFMCNVYELHYCOLDIV 181
 b 125 FADYAPRCPDLKHETTGDALCERGLDDRLILYVTEGFREAVCISNTFLHUGMDKV 184

Qy 182 HINHGDPYVIRPLPFYVOLFMDVNRLVPDPENTHRSIGEGFVYPPFYNTGCHLICHDCV 241
 Db 185 TIGDAVHRIVYPQIQLMFMDFSRVIAIDPNCNERSIGENFTBLPFFNRPLRLFFBAV 244
 Qy 242 JAPMAYALRVBNNTAVARGAHLAFDENHEGALEVPPDITYTYFPOSSSSCTTARGARND 301
 Db 245 VGPAAVALRNRYDVARAARAHLFDENHEGAALPADITPAFEASOG-KPQRGAR-- 299
 302 VNSTKPSPGCFERLASMAAATLAHEVIIFTGIVETBDIDKEVNFMFINEGTLPR 361
 Db 300 --DAGNKGPGAGFCQRLASVMAQDAALAESISMSAWEDEPPPTTPLLQEQTPAIR 357
 Qy 362 INALGSYTAAVGITGAMYFSPNSALYLTBEVEDGMTEAKDGGPSPSNRFYKOFAGPHLA 421
 Db 358 AGAVARYLAKRAGYQAMTSNALSHTEDVDPGPKDKSK-PSYTFPLVGHVA 416
 Qy 422 ANPQTDRDGHVL-----SSQSTGSSNTTEFSYDYLALICGFGAPLIARLLFLYLERCDAG 474
 Db 417 ANPQLDREGVPGYEGRPAPLWGQTQ-BFAGHHLAMLCGSPALLAKMLFLYLERCDGG 475
 Qy 475 AFTGGHG-DALKYUHG-TGTDSEBIPOLSLCKHTPWCANTVHRLQRQMPREGATRQIGN 533
 Db 476 VIVGRQEMDVFYADSGQDGTDPVNLCPTETRACHTLMLMRARHFKFASARGAIGV 535
 Qy 534 PGTMNSQYSPDODPGNAYAPLILRKPGDOTEAKATMDTYRATERLRLPIDEQBERLIDR 593
 Db 536 PGTMNSAYSQCDVGLGNYAAFSAKR-AUGSNTPTTQETYAAERVWMAEJALQYDQ 594
 Qy 594 GAPCESEGJSSV1D�HTPFRILLDTARIEQTTOFMKVLYVETDYKIRESELSEATHSM 653
 Db 595 AVPTALGRLETIGNEREALHTVNNIKOLVDRVEQLMRNLLBGRNKFDRGLAENHAM 654
 Qy 654 ALTFDPSGAFCPITNFLYKTHLAVVODLALSQCHCYFGGCOVEGENFRNCFOPYLRR 713
 Db 655 SLSLDPYTCPCPCLPOLLARRNSIAVYDLSQCHGVPAGGVEGANFRNGFOPVYLR 714
 Qy 714 FVDLFLNGGFISTSTRITVTLSEG-PVSADNPPTGODAPAGRTPDGLARVSVEVDIRVK 772
 Db 715 VMDLNGNGFASKLTATVSEGAICASLTCATQAPLESSEGDYARVTFQPKERVK 774
 Qy 773 NRVYPSGNCTNLSEAAARLVGLASAYQOEKRVDMLRHALGFLQFHGLPFRGMPPN 832
 Db 775 SRVLFGASANASAAKARVASQSOAYQKPDKVDTLGLPQFLQFHAFVFPNKPQPG 834
 Qy 833 SKSENQPWQIWTLLRQNMPADLTUEETTIAAVKRTEEYAINFINLPPCIGBLAQF 892
 Db 835 SNOONPQWWTALQRNQLPARLISREDIETIFIKRPSLDYCAINFNLANNVSEELAMY 894
 Qy 893 YMANLILKYCDHSOYLINTLTSITGARRPDPSSVHLWIRDVTSADITQAKALLEK 952
 Db 895 YMANQILRYCDHSTYFTINTLTAVIAGSRPSPVQAAWAPD--GGAGLEAGARALMDS 951
 Qy 953 TENLPELWTTAFTSTHLYRAAMNORPMVYGLSISISKYHGAAGNNRVEQAGNWGLNGKKN 1012
 Db 952 LDAHFGAWTSMSPASCLNLRPMARPMVGLSISISKYGMANDRVEQAGNWASLGGKKN 1011
 Qy 1013 VCPLEPTDTRTRPLIACPRGGFCPVCPTGPGSSGNTTISDQVRGLIVSGGAMVQLAIYAT 1072
 Db 1012 ACPLLIFDTRKEVLACPAGFVCAASLLGGGAHEHSLCEQRTGILAEGGAAVASSVFA 1071
 Qy 1073 VTRAVGARAQHMAPDDWLSLTDEFLARDLELHDQIQTLTLPWVEGAL---BAVKI 1128
 Db 1072 TVKSLGPRTQQLQEDNLLALDEYLSSEMFETRALEGHEWNTDAALEVHAEAL 1131
 Qy 1129 LDEKTIAQDGETPTLNANFDCEPSHDTTSVNLNISNSNISSTVPCLKRPEPDBELFD 1188
 Db 1132 VSQLGAGE----VFNFGDFGEDDHAAFSFGGLAA--AGAAGVARKRAFHGDPPG 1182
 Qy 1189 LSQIPKHGNITMEM 1203
 Db 1193 -EGGPERK--DLTLDM 1195

Qy	714	FVDLNGGPGLSTRSITUTSEG- PVSAPNPTGQDADAGRTFDGDLARVSVEVDIRVK	772
Db	715	VMDLFNGFLSAKLTVALSEGAIACAPSLTQAGOTAAESSEFGDYARVTLGFPEKLVRK	774
Qy	773	NRVTFSGNTNLSESLAARLVGLASAYAQREQRDVMDLHAGLFLQFHGLPFGRMPPN	832
Db	775	SRVLFAGASANASAAKARVASQYQKPDREVDTLILGLPGLFLLQFQHAFVIFPGKPPG	834
Qy	833	SKSPNPQWTWTLQRNQMPADKLTHEETTAIAKRFTEYYAIAINFNLNPLTCIGELAQF	892
Db	835	SNQNPNPWTALQRNQLPARLISREDIETAFIRKFLSDLGAINFINLNPNTSELAMY	894
Qy	893	YMANLILKYCDHSQYLINTLTSITIIGARRPRDPSSVLHWIRKDVTSAADIETOAKALLEK	952
Db	895	YMANQILRYCDHSTTYFNTLIAVIAGRRPSPVQAIAAWAQO--GGAGLEAGAALMDS	951
Qy	953	TENIPELWTAFTSHLYRAAMNQRPKVYLTGTSISKYHGAGNNRVPQGENWSGLNGKN	1012
Db	952	LDAIPGANWISMFSCLNLRPMVAAPRPMVGLSISXKGAGNDIVFQAGNWSLGGKN	1011
Qy	1013	VCPLFEDTRTRRPTIACPRGGFICPTGPSSGNRRTTSDOVRGIVSGAMVQLAIYAT	1072
Db	1012	ACPLLIFDTRTRKVLACPRAGFCAASSLGGAAHESLCEQLRGIAEGGAAVASVYFA	1071
Qy	1073	VRVAGCAGAQMADFDDWLSLTDEFALRDLEBLHDIIQJLETPTVSEGAL---BAVKI	1128
Db	1072	TVSLGSPRQQIWDMLALEDEVSEUNMFTTRALERGHGEWSTDAALEYAHAEAL	1131
Qy	1129	LDEKKTAGDGETTPNLAFNPDOSCEPHDTSNSVNTGSNTGSNTVPGLKRPPEDELFD	1188
Db	1132	VSQLGAGE-----VNFDGDFGDEDHAASFGGLAAA - AGAGAVKRAFHGDDFG	1182
Qy	1189	LSGTPIKGNTMEM 1203	
Db	1183	-ESPPEEK- DLTDIM 1195	
RESULT 5			
US-10-237-551-231			
; Sequence 231, Application US/10237551			
; Publication No. US0030165820A1			
; GENERAL INFORMATION:			
; APPLICANT: Day, Craig H.			
; APPLICANT: Hosken, Nancy A.			
; APPLICANT: Parsons, Joseph M.			
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND			
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION			
; FILE REFERENCE: 210121_538C3			
; CURRENT APPLICATION NUMBER: US/10/237,551			
; CURRENT FILING DATE: 2002-09-06			
; NUMBER OF SEQ ID NOS: 254			
; SOFTWARE: FastSEQ for Windows Version 4.0			
; SEQ ID NO: 231			
; LENGTH: 274			
; TYPE: PRT			
; ORGANISM: Herpes simplex virus			
; US-10-237-551-231			
Qy	28	LEBISPLAATSDSIALPLMRNLTVEKTFSTSSLAIVSGARTGAGGITLKTTSHF	87
Db	2	LELLSULLSARSGDAVAYAVPLIVLTVESGFEANAVAVGSRRTGCGTAAISLKLPSHY	61
Qy	88	YPSVTVPHGGRHVLLSSAAPNLTCNAARERFGSRQGPYDGAETTGAEICTRLGL	147
Db	62	SPSVYVPHGGRHLAESTQPNLTRLCEARRHGPSDYAPRPODLKHETTGDAICBLRLGL	121
Qy	148	EPENTILYVTAFLKEAFMCNVPLHYGLD1VHTNGDVTRIPFPVOLFMFDYRNRLV	207

Db 680 TALLHSGTAERVPRAADTGRD-AEPAYARVAASHQSVOAVGRPA 724

RESULT 8

US-09-905-129-2
 Sequence 2, Application US/0905129
 Patent No. US20030137705A1
 GENERAL INFORMATION:
 APPLICANT: Einat, et al
 TITLE OF INVENTION: GENES ASSOCIATED WITH MECHANICAL STRESS, EXPRESSION PRODUCTS THEREOF
 FILE REFERENCE: 540579-2007.2
 CURRENT APPLICATION NUMBER: US/09/905,129
 CURRENT FILING DATE: 2001-07-13
 PRIOR APPLICATION NUMBER: US/09/802,318
 PRIOR FILING DATE: 2001-03-08
 PRIOR APPLICATION NUMBER: 60/207,821
 PRIOR FILING DATE: 2000-05-30
 PRIOR APPLICATION NUMBER: 60/084,044
 PRIOR FILING DATE: 1998-05-11
 PRIOR APPLICATION NUMBER: 60/085,673
 NUMBER OF SEQ ID NOS: 25
 SEQ ID NO 2
 SOFTWARE: Patentin version 3.0

LENGTH: 2597
 TYPE: PRT
 ORGANISM: *rattus species*

FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (1)-(2597)
 OTHER INFORMATION: 'x' can be any amino acid

US-09-905-129-2

Query Match 1.9%; Score 120.5; DB 10; Length 2597;
 Best Local Similarity 1.8%; Pred. No. 0.46; Mismatches 113; Gaps 40;
 Matches 145; Conservative 113; Mismatches 287; Indels 233;

2Y 416 AGPHLAANPQID--RGDHVLUSSGTGSSNTE-----FSDVYALICG 455
 2Y 274 SGAFLCTKPTIDPSKSFKSLTQEONGSASTSPQDFIEPGSLSLNMIXXSGNADWYC 333
 2Y 456 FGAPLLARLFLYLERCDAGFTGGHDALKYVTGTFDSEIPCSLCEKTRPV--CAHT 511
 2Y 334 IQKP-----SRTSPTAFTENDYIM-LNASFSTNQVSYSDCPLGNATAPILLRKGD 561
 2Y 512 TVHLRQORMP-----REGQATRQPIGFEGTMNSQYSDCPLGNATAPILLRKGD 561
 2Y 383 DSPLIWERKPOLTEPSLSSRYKQVALPRDPEDFISLEADVR-APPFWQEQEKVLQNLRT 441
 2Y 562 QTEAAKATMDTYATLERFLIDQERL-----LDR-----GAPCSSBG 601
 2Y 442 ATTLSLTIQIQSTDQIAQALPRAEERKWTMILMMNPKEPLTVLNGGTLAISCPKG 501
 2Y 602 ----LSSVIVDHPFR-----RILDITLRAE-QNTTOFMKV-----VETRD---- 639
 2Y 502 DPSPHLEWLADGSKVRAPYSEDRGRLIDKNGCLEQMDADSFAGLYHCISTNDADAV 561
 2Y 640 --YKIR--EGLSEATHSMALFDPSGAF---CPITNLFVKEFHLYQDIALSQCHV 691
 2Y 562 LTYRITVVEPYGESTDGYQHTVVTGETLDLPLSTG-----VPDASIS--WI 608
 2Y 692 FYGGQVEGRNFRQFQPVTLRRRFYDLENGGFISRSRITYLSEGPPVS--APNPT---- 743
 2Y 609 LPGNNTVFSQPSRDR-----QILANGTQLRQLQT-----VPDASIS--WI 608
 2Y 744 -----LGDDAPAGRTRFGDALARVSVEIRDTRVQRVVFQGNCNTNLSEARAR 791
 2Y 658 PRVSVQRKGGRMVAHDREAGSGIGE-PNSSVSKLQPSLK---LSASALTGSBAGKQ- 711
 2Y 792 LVGLASAYQRQEKRVDMLGALG-FLLKQFHGILFPRMPPNSKSPNPOWFMLQ--R 847
 b 712 ---VSGVHERKRNKRDLTHRRGJSTLRFRE--HRLQPLSARRIDPQRMAALLBKAKK 765

Qy 848 NQMPADKLTHEEITT-----IAAVKRTEEEVAAINFINLPPTCIGBLAQFYMANLILK 900
 Db 766 NSVP----KQENTTIVKPVFLAYLVELTDEERDASGMI--PP----- 802
 Qy 901 YCDHSQYLNLTSLTGARRPRDPSSYLHWIKRDVTSAAIDETOQALLEKTNLPBLW 960
 Db 803 --DEFMFVLTAKASGVGRSPPTADSGPNTFHGMNTIASGTEVSTVNPFQTLQ-SEHLDFK 859
 Qy 961 TTAFSTHIVRAAM-----NORMPVYGLISTSKVHGAAGNNRVQAGNMSG-- 1006
 Db 860 LESVTINGTAVTKSNPSTASKIEDTTNPNIIFFP-SVAEIRDSA----CAGRASSQS 912
 Qy 1007 ---LINGKRNVCPLFTEDTRRPIACPRGGFI-----CPVTG--PSSGN 1045
 Db 913 AHPVYGGN---NATYGHNTYSSFTSKASTVTLQPINPTESYGPQIPITGVSRPSSD 966

RESULT 9

US-09-905-129-10
 Sequence 10, Application US/0905129
 Patent No. US20030137705A1
 GENERAL INFORMATION:
 APPLICANT: Einat, et al
 TITLE OF INVENTION: GENES ASSOCIATED WITH MECHANICAL STRESS, EXPRESSION PRODUCTS THEREOF
 FILE REFERENCE: 540579-2007.2
 CURRENT APPLICATION NUMBER: US/09/905,129
 CURRENT FILING DATE: 2001-07-13
 PRIOR APPLICATION NUMBER: 09/802,318
 PRIOR FILING DATE: 2001-03-08
 PRIOR APPLICATION NUMBER: 60/207,821
 PRIOR FILING DATE: 2000-05-30
 PRIOR APPLICATION NUMBER: 60/084,944
 PRIOR FILING DATE: 1998-05-11
 PRIOR APPLICATION NUMBER: 60/085,673
 PRIOR FILING DATE: 1998-05-15
 NUMBER OF SEQ ID NOS: 25
 SOFTWARE: Patentin version 3.0
 SEQ ID NO 10
 LENGTH: 2597
 TYPE: PRT
 ORGANISM: *Rattus sp.*
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (1)-(2597)
 OTHER INFORMATION: 'x' can be any amino acid

US-09-905-129-10

Query Match 1.9%; Score 120.5; DB 10; Length 2597;
 Best Local Similarity 1.8%; Pred. No. 0.46; Mismatches 113; Gaps 40;
 Matches 145; Conservative 113; Mismatches 287; Indels 233;

Qy 416 AGPHLAANPQID--RDGHVLUSSGTGSSNTE-----FSDVYALICG 455
 Db 274 SGAFLCTKPTIDPSKSFKSLTQEONGSASTSPQDFIEPGSLSLNMIXXSGNADWYC 333
 Qy 456 FGAPLLARLFLYLERCDAGFTGGHDALKYVTGTFDSEIPCSLCEKTRPV--CAHT 511
 Db 334 IQKP-----SRTSPTAFTENDYIM-LNASFSTNQVSYSDCPLGNATAPILLRKGD 561
 Qy 512 TVHLRQORMP-----REGQATRQPIGFEGTMNSQYSDCPLGNATAPILLRKGD 561
 Db 383 DSPLIWERKPOLTEPSLSSRYKQVALPRDPEDFISLEADVR-APPFWQEQEKVLQNLRT 441
 Qy 562 QTEAAKATMDTYATLERFLIDQERL-----LDR-----GAPCSSBG 601
 Db 602 ----LSSVIVDHPFR-----RILDITLRAE-QNTTOFMKV-----VETRD---- 639
 Db 502 DPSPHLEWLADGSKVRAPYSEDRGRLIDKNGCLEQMDADSFAGLYHCISTNDADAV 561
 Qy 640 --YKIR--EGLSEATHSMALFDPSGAF---CPITNLFVKEFHLYQDIALSQCHV 691
 Db 562 LTYRITVVEPYGESTDGYQHTVVTGETLDLPLSTG-----VPDASIS--WI 608
 Qy 692 FYGGQVEGRNFRQFQPVTLRRRFYDLENGGFISRSRITYLSEGPPVS--APNPT---- 743
 Db 609 LPGNNTVFSQPSRDR-----QILANGTQLRQLQT-----VPDASIS--WI 608
 Qy 744 -----LGDDAPAGRTRFGDALARVSVEIRDTRVQRVVFQGNCNTNLSEARAR 791
 Db 442 ATTLSLTIQIQSTDQIAQALPRAEERKWTMILMMNPKEPLTVLNGGTLAISCPKG 501
 Qy 658 PRVSVQRKGGRMVAHDREAGSGIGE-PNSSVSKLQPSLK---LSASALTGSBAGKQ- 711
 Qy 792 LVGLASAYQRQEKRVDMLGALG-FLLKQFHGILFPRMPPNSKSPNPOWFMLQ--R 847
 Db 803 --DEFMFVLTAKASGVGRSPPTADSGPNTFHGMNTIASGTEVSTVNPFQTLQ-SEHLDFK 859
 Qy 860 LESVTINGTAVTKSNPSTASKIEDTTNPNIIFFP-SVAEIRDSA----CAGRASSQS 912
 Db 901 YCDHSQYLNLTSLTGARRPRDPSSYLHWIKRDVTSAAIDETOQALLEKTNLPBLW 960
 Qy 961 TTAFSTHIVRAAM-----NORMPVYGLISTSKVHGAAGNNRVQAGNMSG-- 1006
 Db 1007 ---LINGKRNVCPLFTEDTRRPIACPRGGFI-----CPVTG--PSSGN 1045
 Db 1014 AHPVYGGN---NATYGHNTYSSFTSKASTVTLQPINPTESYGPQIPITGVSRPSSD 966

2Y 640 -YKIR--EGLSEATHSMALTFDDPSGAF---CPITNFLVKRTHLAVYQDLALSQCHCV 691
 Jb 562 LTVRTIVVEPYGESEHDGVQHTVVVGETDLPLCSTG-----VPDASIS---WI 608
 Y 692 FYGOOVEGERFRNQQPVLRFFDLENGFIFSRSTIVTLLSGPVS -APNPT---- 743
 Jb 609 LPNTNTVFSQPSDR-----QILANGTLRLQVT PKDQGHYOCVAANPSGADFSS 657
 Y 744 -----LGODAPAGRFTFDGLARVSVEIRD-PRVNRVFSCTNC-SEARAR 791
 Jb 658 FKVSVQKGGRMVRHDEAGGGIGE -PNSSVSKQPSLK---ISASALTGEASKQ- 711
 Y 792 LVGLASAYQREKRVDMLELG -FLLOFGRGLI PGRMNPPNSKSPNQFWTLLQ--R 847
 Jb 712 ---VSGVERKRNKRDLTERRGDSLRRPRE --HROPLPSARIDQRWALLEKAKK 765
 Y 848 NQWNAFKLTHEETT-----IAVAKRFTTEAYAINFINLPCTIGELAQFYMANJLK 900
 Jb 766 NSVP---KKQENTNTVKVPLAYFLVELTDEKDASGM--DP----- 802
 Y 901 YCDHSQYLINTLTSITGARRPRDPSVHLWIRKDVTSAADIETOAKALLEKTNPELM 960
 Jb 803 -DEFMVKTRASGVPSRSPADISGPVNQGMNTSISGTEVNTNPOTLQ-SEHLPDFX 859
 Y 961 TTATPTSTHLVRAAM-----NORPMVTLGIGISKYHGAGNNRYFQAGNWSG-- 1006
 Jb 860 LFSVNTGTAVTNSMNPPIASKIEDTTQNPIIIIFP-SVAEIRDSA-----QAGRSSQS 912
 Y 1007 ---LNGGRNVCPLFTEDETRRIACPRGGF----CPVTG---PSSGN 1045
 Jb 913 AHPTVGGN---MATYGHNTNTSSFTSKASTVLQPINPTESYGPQIPITGVSRPSSSD 966

JRESULT 10
 ;-09-905-129-13
 Sequence 13, Application US/0905129
 Patent No. US200902037705A1 /US/0905129
 GENERAL INFORMATION:
 APPLICANT: Elnat, et al
 TITLE OF INVENTION: GENES ASSOCIATED WITH MECHANICAL STRESS, EXPRESSION PRODUCTS THERE
 TITLE OF INVENTION: AND USES THEREOF
 FILE REFERENCE: 540579-2007.
 CURRENT APPLICATION NUMBER: US/09/905,129
 PRIOR APPLICATION NUMBER: 09/802,318
 PRIOR FILING DATE: 2001-07-13
 PRIOR APPLICATION NUMBER: 09/802,318
 PRIOR FILING DATE: 2001-03-08
 PRIOR FILING DATE: 2000-05-30
 PRIOR FILING DATE: 1998-05-11
 PRIOR FILING DATE: 1998-05-15
 NUMBER OF SEQ ID NOS: 25
 LENGTH: 2597
 TYPE: PRT
 ORGANISM: Rattus sp.
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (1) _ (2597)
 OTHER INFORMATION: 'x' can be any amino acid
 ;-09-905-129-13

Query Match 1.9%; Score 120.5; DB 10; Length 2597;
 Best Local Similarity 18.6%; Pred. No. 0..46; Mismatches 113; MisMatches 287; Index 233; Gaps 40;
 SEQ ID NO: 13
 LENGTH: 2597
 TYPE: PRT
 ORGANISM: Rattus species
 FEATURE:
 NAME/KEY: misc_feature

RESULTS 11
 Sequence 2, Application US/09991630
 Patent No. US20151514A1
 GENERAL INFORMATION:
 APPLICANT: Elnat, et al
 TITLE OF INVENTION: GENES ASSOCIATED WITH MECHANICAL STRESS, EXPRESSION PRODUCTS THERE
 FILE REFERENCE: 540579-2007.3
 CURRENT FILING DATE: 2001-11-06
 PRIOR APPLICATION NUMBER: 09/905,129
 PRIOR FILING DATE: 2001-07-13
 PRIOR APPLICATION NUMBER: 09/802,318
 PRIOR FILING DATE: 2001-03-08
 PRIOR APPLICATION NUMBER: 09/729,485
 PRIOR FILING DATE: 2000-12-04
 NUMBER OF SEQ ID NOS: 28
 LENGTH: 2597
 TYPE: PRT
 ORGANISM: Rattus species
 FEATURE:
 NAME/KEY: misc_feature

416 AGPHIAANPDT---RDGHVLLSSQTSNTE-----FSYDYLICG 455
 274 SGAFICTKPTIDPSIKSLVQEDNGSAASTSPDIFBFGSISLNLNTXXSGNKDNKVC 333
 456 FGAPPARIYLFLYERCDAGAFTGGHGDALKYVFTDFSEIPCSLCXHTRPV---CAHT 511

Db 334 IQKP-----SRTSPARTBEENDYM--LNASFSTNLVCSVDNHITQPVWOLLALYS 382
 Qy 512 TVRHRQMRP-----RGFOATRQFPGFETMNSQSYSDCDPLGNYAPYLIRKGD 561
 Db 383 DSPLLERKEQLTETPSLSRKYKQVALPBDIFTSIAEVR-ADPFWFQQEIKVLOQNRT 441
 Qy 562 OTEAAKATMDTYRATERLFLDLEQERU-----LDR-----GAPCSSEB 601
 Db 442 ATLSLQLQDPSITAQIALPRAENRAERIKWTMNLMMNNPKLERTVLYGGT-ALSCPKG 501
 Qy 602 -----LSSVIVDHPPTF----RILDILRARIE-QTTQMVKL--VETRD---- 639
 Db 502 DPSPHLEWILADGSKVRAPYVSEDGRILDKNCNLQMASSDAGLYHCISNDADAV 561
 Qy 640 -YKIR-ECUUSETHSMALTFDYSGAF---CPINFLVRIRTHLAVVODALSQCHCV 691
 Db 562 LYTR-TVVEPYGESTHDSSVQHTVVTGSTDLDLPCLSIG-----VPDYSIS--WI 608
 Qy 692 FYGOOVBGNFRNFRNQPVLRRFVLDLNGFFISTRSTRVTLSEGPPVS--APNPT---- 743
 Db 609 LGPNVTVFSESPRBR-----QIANGTRILQVTPQKGHQCVAAAPSGADFSS 657
 Qy 744 -----LGQDADAPGRATFDGLARVSTEVIRDTRVKNRVRVSGNCNTLSEEARAR 791
 Db 658 FKVSVQKGQRMVYEHDRBAGGSGLGE-PNSSYSLKQPSLKL---LSASALTGSEAGQ- 711
 Qy 792 LVGLASAYQREKVDMLHGAJG-FLIKQFQGFLERGMPPNKSNSDNPWTLLC--R 847
 Db 712 ---VSGVERKNGHDLIHRRGDSLURRFRE--HEROLPLSARRDPORWALLEKAKK 765
 Qy 848 NQMPADKLTHEETT-----IAAVKRFTTEYYAANFINLPTCTIGELAQFYMANLILK 900
 Db 766 NSVP---KKQENTNTVKVPLAYFLVELTDEKDASGM--DP----- 802
 Qy 901 YCDHSQYLINTLTSITGARRPRDPSVHLTRKDVTSAADIETOAKALLEKTNPELM 960
 Db 803 -DEEFMVWLKTAKGVPGRSPADSPVNHGFMTSIASGTEVSTNPQTQ-SEHLPDFK 859
 Qy 961 TRAFTSTHLVRAAM-----NORPMVVLGIGISKYHGAGNNRYFQAGHWSG-- 1006
 Db 860 LFSVNTGTAVTNSMNPPIASKIEDTTQNPIIIIFP-SVAEIRDSA-----QAGRSSQS 912
 Qy 1007 ---LNGGRNVCPLFTEDETRRIACPRGGF----CPVTG---PSSGN 1045
 Db 913 AHPTVGGN---MATYGHNTNTSSFTSKASTVLQPINPTESYGPQIPITGVSRPSSSD 966.

LOCATION: (1)..(2597)
 OTHER INFORMATION: 'x' can be any amino acid
 Query Match 1.9%; Score 120.5; DB 10; Length 2597;
 Best Local Similarity 18.6%; Pred. No. 0.46; Matches 113; Gaps 40;
 Matches 145; Conservative 113; Mismatches 287; Indels 233; Gaps 40;
 416 AGPHLAANPQTD--RDGHVILSSQSSTSSNTE-----FSDYDIALIGC 455
 274 SGAFLCTKPTPSLKSLSLVIQEDNSAASASTSPQDFIEBPFGSLSIANTXXSGNKADMVCs 333
 456 FAAPILLARLLFLVERCDAGAFGFGHDAKVVGTGFDFSEIPSLCLCEKHTRPV---CAHT 511
 1QKP-----SRTSPTAAFEENDYIM-LNAFESTNLCSVDYHNIHQPVWQLIALYS 382
 512 TVHLRQMRP-----REFGQATRQPPIQVFGTNNKSOYSDCDPLGNAYAPYLILRKPGD 561
 383 DSPLILERKPQIETPLASSRYKQVALRPEIDFTSIEADVR-ADPFWFQQEKIVIQLNRT 441
 562 QTEAAKATMDDTYRATLERLFDLEQRL-----LDR-----GACPSSEG 601
 442 ATTLSLTIQIQFSTDAQIALPRAEMRAERLKTMILMNPNPLERTLVGGTIALSCP GK 501
 602 ---LSSVIDDHPTF-----RILDTRBARIE OTTFQFMVL---VETRD---- 639
 502 DPSPHLWLADGSKVRAPIVYSEDGRLLIDDKLQMAQDSEFDAGLYHCISTNDADADV 561
 640 --YKTR-EGGEATHSMALTFDPYSGAF---CPTINFLFLYKTRTHLAWVODALSQCHCV 691
 562 IITYRIVVEPPIGESTDGCVHTVVTGETLILPCSTG-----VDAStIS--WI 608
 692 FYQQOQEGRNRFRNQOFQPVLRRFVDLFNGGPISTRSITVILSEGPVS--APNPT----
 609 LPGNTVFSQPSRDR-----QIINNGTIRLQYT-PKDQGHYQCVVANPSGADESS 657
 744 -----LGQDAPAGTEDGDLARYSEVTRDIRYKRNVVFSGNTNLSEARRAR 791
 658 FKVSVQKGQMRVHDREAGGSSLGE-PNSSVLQFQPALSLK---LSASALTGEAKQ-
 792 LVGLASAYQRDEBKRVDMHLGALG-FILKQFHGLLFRGMMPNSKSPNPWFVTLQ---R 847
 712 ---VSGVHRKNCNKHDLIIRRGDSTLRRFRE-HRQLPLISARRIDPQWALLEKAK 765
 848 NQMPADRLTHEITT-----IAAVKRFTEEYAAINFNMUPTCIGELAQFYMANLILK 900
 766 NSVP---KQDENTVVKPVPLAVPLVLTDEKDASGMI--PP----- 802
 901 YCDHSQYLINMLTSITITGARRPRDSSVLEWIRKDTSAADIEOKALLEKTNLPEW 960
 803 -DEEFMVWIKTAKSGPGRSTADSGPVGNGFMUTIASGSEVSTYNPQLO-SHLPDFX 859
 961 TIAFTSTHLYRAM-----NQRMVWGLISITISKYGAAGNNRVRQAGNWSG- 1006
 860 LFSVINGTAVTKSMNPPIASKIEDTQNPIIIFP-SVARIRDSA-----QAGRASSOS 912
 1007 ---LNQGKNVYCPLFEDRTRFLIACPRGFFI-----CPYTG--PSSGN 1045
 913 AHPVITGGN--MATIGHNTNYYSSFTSKASTVLPQINPTESYGQPIPTGVSRSSD 966
 US-09-991-630-10
 SUIT 12
 -0-991-630-10
 Sequence 10, Application US/09991630
 Patent No. US20151514A1
 GENERAL INFORMATION:
 APPLICANT: Binat, et al.
 TITLE OF INVENTION: GENES ASSOCIATED WITH MECHANICAL STRESS, EXPRESSION PRODUCTS THEREOF
 FILE REFERENCE: 54057-2007.3
 CURRENT APPLICATION NUMBER: US/09/991, 630
 CURRENT FILING DATE: 2001-11-06
 PRIORITY DATE: 2001-07-13
 PRIOR APPLICATION NUMBER: 09/802, 318
 PRIOR FILING DATE: 2001-03-08
 PRIOR APPLICATION NUMBER: 09/729, 485
 PRIOR FILING DATE: 2000-12-04
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO: 10
 LENGTH: 2597
 TYPE: PT
 ORGANISM: Rattus species
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (1)..(597)
 OTHER INFORMATION: 'x' can be any amino acid
 US-09-991-630-10
 Query Match 1.9%; Score 120.5; DB 10; Length 2597;
 Best Local Similarity 18.6%; Pred. No. 0.46; Matches 113; Mismatches 287; Indels 233; Gaps 40;
 Matches 145; Conservative 113; Mismatches 287; Indels 233; Gaps 40;
 416 AGPHLAANPQTD--RDGHVILSSQSSTSSNTE-----FSDYDIALIGC 455
 Qy 274 SGAFLCTKPTPSLKSLSLVIQEDNSAASASTSPQDFIEBPFGSLSIANTXXSGNKADMVCs 333
 Db 456 FGAPIPLARLLFLYERCDAGAFGFGHDAKVVGTGFDFBIBPCSLICEKFTRPV---CAHT 511
 Db 334 IQKP-----SRTSPTAAFEENDYIM-LNASFSTNLCSVDYHNIHQPVWQLIALYS 382
 Qy 512 TVHLRQMRP-----RFGATRQFTRGIVGTMNQYQSDCDPLGNAYAPYLILRKGD 561
 Db 383 DSPLILERKPQIETPLSSRYKOVALRPEIDFTSIEADVR-ADPFWFQKEVTLQMR 441
 Qy 562 QTEAAKATMDDTYRATLERLFDLEQRL-----LDR-----GAPCSSEG 601
 Db 442 ATTLSLTIQIQFSTDAQIALPRAEMRAERLKTMILMNPNPLERTLVGGTIALSCP RG 501
 Qy 602 -----LSSVTVDHPTR-----RILDTRBARIE OTTFQFMVL---VETRD---- 639
 Db 502 DPSPHLWLADGSKVRAPIVYSEDGRLLIDDKLQMAQDSEFDAGLYHCISTNDADADV 561
 Qy 640 --YKTR-EGGEATHSMALTFDPYSGAF---CPTINFLFLYKTRTHLAVVODALSQCHCV 691
 Db 562 LTRITIVTPYGCSTHDSGVQHVTVTGETLIDPLCISIG-----VPDASIS--WI 608
 Qy 692 FYQQOQEGRNRFRNQOFQPVLRRFVDLFNGGPISTRSITVILSEGPVS--APNPT----
 Db 744 -----LGQDAPAGTEDGDLARYSEVTRDIRYKRNVVFSGNTNLSEARRAR 791
 Qy 752 FKVSVQKGQMRVHDREAGGSSLGE-PNSSVLQFQPALSLK---LSASALTGEAKQ-
 Db 792 LVGLASAYQRDEBKRVDMHLGALG-FILKQFHGLLFRGMMPNSKSPNPWFVTLQ---R 847
 Qy 848 NQMPADRLTHEITT-----IAAVKRFTEEYAAINFNMUPTCIGELAQFYMANLILK 900
 Qy 860 LFSVINGTAVTKSMNPPIASKIEDTQNPIIIFP-SVARIRDSA-----QAGRASSOS 912
 Db 901 YCDHSQYLINMLTSITITGARRPRDSSVLEWIRKDTSAADIEOKALLEKTNLPEW 960
 Qy 903 -DEEFMVWIKTAKSGPGRSTADSGPVGNGFMUTIASGSEVSTYNPQLO-SHLPDFX 859
 Db 961 TIAFTSTHLYRAM-----NQRMVWGLISITISKYGAAGNNRVRQAGNWSG- 1006
 Qy 966 NSTP---KKQENTTIVKPVPLAVPLVLTDEKDASMI--PP----- 802
 Qy 901 YCDHSQYLINTLTSITITGARRPRDSSVLEWIRKDTSAADIEOKALLEKTNLPEW 960
 Db 803 -DEEFMVWIKTAKSGPGRSTADSGPVGNGFMUTIASGSEVSTYNPQLO-SHLPDFX 859
 Qy 961 TIAFTSTHLYRAM-----NQRMVWGLISITISKYGAAGNNRVRQAGNWSG- 1006
 Db 1007 ---LINGKNCVCPLFEDRTRFLIACPRGFFI-----CPVIG--PSSGN 1045
 Qy 913 AHPVITGGN--MATIGHNTNYYSSFTSKASTVLPQINPTESYGQPIPTGVSRSSD 966
 US-09-991-630-10
 SUIT 12
 -0-991-630-10
 Sequence 10, Application US/09991630
 Patent No. US20151514A1
 GENERAL INFORMATION:
 APPLICANT: Binat, et al.
 TITLE OF INVENTION: GENES ASSOCIATED WITH MECHANICAL STRESS, EXPRESSION PRODUCTS THEREOF
 FILE REFERENCE: 54057-2007.3
 CURRENT APPLICATION NUMBER: US/09/991, 630
 CURRENT FILING DATE: 2001-11-06
 PRIORITY DATE: 2001-07-13
 PRIOR APPLICATION NUMBER: 09/802, 318
 PRIOR FILING DATE: 2001-03-08
 PRIOR APPLICATION NUMBER: 09/729, 485
 PRIOR FILING DATE: 2000-12-04
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO: 10
 LENGTH: 2597
 TYPE: PT
 ORGANISM: Rattus species
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (1)..(597)
 OTHER INFORMATION: 'x' can be any amino acid
 US-09-991-630-10
 Query Match 1.9%; Score 120.5; DB 10; Length 2597;
 Best Local Similarity 18.6%; Pred. No. 0.46; Matches 113; Mismatches 287; Indels 233; Gaps 40;
 Matches 145; Conservative 113; Mismatches 287; Indels 233; Gaps 40;
 416 AGPHLAANPQTD--RDGHVILSSQSSTSSNTE-----FSDYDIALIGC 455
 Qy 274 SGAFLCTKPTPSLKSLSLVIQEDNSAASASTSPQDFIEBPFGSLSIANTXXSGNKADMVCs 333
 Db 456 FGAPIPLARLLFLYERCDAGAFGFGHDAKVVGTGFDFBIBPCSLICEKFTRPV---CAHT 511
 Db 334 IQKP-----SRTSPTAAFEENDYIM-LNASFSTNLCSVDYHNIHQPVWQLIALYS 382
 Qy 512 TVHLRQMRP-----RFGATRQFTRGIVGTMNQYQSDCDPLGNAYAPYLILRKGD 561
 Db 383 DSPLILERKPQIETPLSSRYKOVALRPEIDFTSIEADVR-ADPFWFQKEVTLQMR 441
 Qy 562 QTEAAKATMDDTYRATLERLFDLEQRL-----LDR-----GAPCSSEG 601
 Db 442 ATTLSLTIQIQFSTDAQIALPRAEMRAERLKTMILMNPNPLERTLVGGTIALSCP RG 501
 Qy 602 -----LSSVTVDHPTR-----RILDTRBARIE OTTFQFMVL---VETRD---- 639
 Db 502 DPSPHLWLADGSKVRAPIVYSEDGRLLIDDKLQMAQDSEFDAGLYHCISTNDADADV 561
 Qy 640 --YKTR-EGGEATHSMALTFDPYSGAF---CPTINFLFLYKTRTHLAVVODALSQCHCV 691
 Db 562 LTRITIVTPYGCSTHDSGVQHVTVTGETLIDPLCISIG-----VPDASIS--WI 608
 Qy 692 FYQQOQEGRNRFRNQOFQPVLRRFVDLFNGGPISTRSITVILSEGPVS--APNPT----
 Db 744 -----LGQDAPAGTEDGDLARYSEVTRDIRYKRNVVFSGNTNLSEARRAR 791
 Qy 752 FKVSVQKGQMRVHDREAGGSSLGE-PNSSVLQFQPALSLK---LSASALTGEAKQ-
 Db 792 LVGLASAYQRDEBKRVDMHLGALG-FILKQFHGLLFRGMMPNSKSPNPWFVTLQ---R 847
 Qy 848 NQMPADRLTHEITT-----IAAVKRFTEEYAAINFNMUPTCIGELAQFYMANLILK 900
 Qy 860 LFSVINGTAVTKSMNPPIASKIEDTQNPIIIFP-SVARIRDSA-----QAGRASSOS 912
 Db 901 YCDHSQYLINMLTSITITGARRPRDSSVLEWIRKDTSAADIEOKALLEKTNLPEW 960
 Qy 903 -DEEFMVWIKTAKSGPGRSTADSGPVGNGFMUTIASGSEVSTYNPQLO-SHLPDFX 859
 Db 961 TIAFTSTHLYRAM-----NQRMVWGLISITISKYGAAGNNRVRQAGNWSG- 1006
 Qy 966 NSTP---KKQENTTIVKPVPLAVPLVLTDEKDASMI--PP----- 802
 Qy 901 YCDHSQYLINTLTSITITGARRPRDSSVLEWIRKDTSAADIEOKALLEKTNLPEW 960
 Db 803 -DEEFMVWIKTAKSGPGRSTADSGPVGNGFMUTIASGSEVSTYNPQLO-SHLPDFX 859
 Qy 961 TIAFTSTHLYRAM-----NQRMVWGLISITISKYGAAGNNRVRQAGNWSG- 1006
 Db 1007 ---LINGKNCVCPLFEDRTRFLIACPRGFFI-----CPVIG--PSSGN 1045
 Qy 913 AHPVITGGN--MATIGHNTNYYSSFTSKASTVLPQINPTESYGQPIPTGVSRSSD 966

RESULT 13-91-630-13
 Sequence 13, Application US/09991630
 Patent No. US2002015151A1
 GENERAL INFORMATION:
 APPLICANT: Einat, et al
 TITLE OF INVENTION: GENES ASSOCIATED WITH MECHANICAL STRESS, EXPRESSION PRODUCTS THEREOF, AND USES THEREOF
 FILE REFERENCE: 540579-007-3
 CURRENT APPLICATION NUMBER: US/09/991,630
 CURRENT FILING DATE: 2001-11-06
 PRIOR APPLICATION NUMBER: 09/905,129
 PRIOR FILING DATE: 2001-07-13
 PRIOR APPLICATION NUMBER: 09/802,318
 PRIOR FILING DATE: 2001-03-08
 PRIOR APPLICATION NUMBER: 09/729,485
 PRIOR FILING DATE: 2000-12-04
 NUMBER OF SEQ ID NOS.: 28
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO: 13
 LENGTH: 2597
 TYPE: PRT
 ORGANISM: Rattus species
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (1)..(2597)
 OTHER INFORMATION: 'x' can be any amino acid

Query Match 1.9%; Score 120.5; DB 10; Length 2597;
 Best Local Similarity 18.6%; Prey: No. 0.46; Mismatches 287; Indels 233; Gaps 40;
 Matches 145; Conservative 113; Mismatches 287; Indels 233; Gaps 40;

Y 416 AGPILAAANQTD--RDGHVILSSGSSNTE----FSYDYLICG 455
 b 274 SGAACTCTKTIDPILSKVSLVTOHDNGSASTSPQDFTEPGSLSLNMXTXSNKADMVCS 333
 Y 456 FGAPILLARLLFYLERCDAAGPTGGHDALKYVTGFDSEBIPSSLCEKHTRPV---CAHT 511
 b 334 IQK-----SRTSPPAFTENDYM-LNASFSTNLVCSDVTHIOPWQLLALYS 382
 Y 512 TVHRLRQMP-----RGFGATROPIGVGTNSQYSDCDPLGNYAPYLILRKPGD 561
 b 383 DSPLILERKPKQLETPSLSRKYQVALRPEDIFTSIEADVR ADPFMFOQEKKIVLQLNRT 441
 Y 562 QTEAKATMQDTYRATLFLIEQEL----LDR----GAPQSEG 601
 b 442 ATTISTLQIQFSTAQIALPRAEMLRKLWMMNNPKLERTVILGGTIALSPGKG 501
 Y 602 ----LSSYIVDHPTER-----RILDTLRARI-QTTOFMKVY--VETRD---- 639
 b 502 DPPSHLEWLADGSKVKRAFYVSSDGRILIDKNGKLELQMAADFGLYHC1STNDADAVY 561
 Y 640 --YKIR--EGLESEATHSMALTPDYPSCAF---CPIINFLVKRTHLAVQDQLLSCHCV 691
 b 562 LTVRITVYVEPYGSTHDGVQRTVYGETLDLPCLSIG-----VPDASIS--WI 608
 Y 692 FIGQVEQENGFNQFPVLRREPVDFNGGFFLTSRSTVUSEGPVS--ANPNT---- 743
 b 609 LPNTVFSQPSDR-----QINNGTIRLQVTPDQGHYQCVAAANPSSADFSS 657
 Y 744 -----LGQDAPAGRTEGDGLARSVEVDIRVKNRVRVFSGNCTNLSSAARAR 791
 b 658 FKVAVQKQKQRMNHEHDPDAGGGGE-PNSSTSLKQFASL ---ISAATGSGAQK- 711
 Y 792 LVGLASAYOQEKEVDMHGAQ-FLIKOFHGHLPGMPPNSKSPNPONWTLLQ--R 847
 b 712 ---VSGVERKNGRDLTHRRRDSITRFR--HQRQLPSARRDQPWAALLEKAKK 765
 Y 848 NQMEADKTCTHEET-----IAVAKRFTEFYAAINFNFPTCIGELAQFYMANLKK 900
 b 766 NSVP-----KKQENTTVKPVPLAVPLVLTDEEKDASSMI--PP----- 802

RESULT 14
 US-10-329-079-11
 Sequence 11, Application US/10329079
 Publication No. US20030198981A1
 GENERAL INFORMATION:
 APPLICANT: FARNET, Chris
 APPLICANT: ZAZOPULOUS, Emmanuel
 APPLICANT: STAFFA, Alfredo
 TITLE OF INVENTION: GENES AND PROTEINS INVOLVED IN THE BIOSYNTHESIS OF LIPOPEPTIDES
 FILE REFERENCE: 3002-111US
 CURRENT APPLICATION NUMBER: US/10/329,079
 NUMBER OF SEQ ID NOS: 66
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO: 11
 LENGTH: 5245
 TYPE: PRT
 ORGANISM: Streptomyces fradiae
 US-10-329-079-11

Query Match 1.9%; Score 120; DB 12; Length 5245;
 Best Local Similarity 20.2%; Prey: No. 18; Mismatches 138; Indels 46; Gaps 68;
 Matches 262; Conservative 138; Mismatches 434; Indels 46; Gaps 68;

Y 102 PSSAAPNLTRACNAARERFGFRCQGPP---VDAVETTGABICPCL----GL 147
 Db 464 PAAAEPAADEGLEAVCDPARRQAATPAPAVQGPVLTFAEADAVSRLLRSRG 523
 Y 148 EPENTILLYVTAFLKEAVFMCMVLFHYGGDLVHINGDVRVRIPLFPVQLMFDVNRLV 207
 Db 524 GPE ---VRVACLDRNAWLPVV-----LATVSGATHVPL----- 556
 Y 208 PDPNTNHRSIGEGFVYPTPPNTGGLCHIHCVIAMAVLVRVNTAVGAAHLAF- 266
 Db 557 -DPRSPHER-----LAVERE-VALLVLAERATEAAVADLAAAPVVLV 597
 Y 267 -DENHEGAV-----LPPDITYTYFQSSSC-----TTARG-----A 297
 Db 598 DDPSTEAAIDALDGPVNTDAADTAPLPGHAAVYVHTSGSTORPKGYTVDRHGLSBLQ 657
 Y 298 RRNVNNTSKPSPSG-GFERRASI-----MAADTAHLHA----- 330
 Db 658 HRRVTFSRIRPSACGPRAAHTYSSPSDASWDLLANVAGHLEMIDDEDLPDPGVVAY 717
 Y 331 -----EVIENTGUYE---TT-----DIKEWPMFIGEGTLRPNL 365
 Db 718 FDRERIDYVDTLPTYFRSLLDAGLLEGPFCISLVALGGEMDGEWLRRAAPRVTM 777
 Y 366 GSY----TARVA-----GVIGAMVFSSENSAAYLVEDSGMTEAKDGGPSFNR 11
 Db 778 NTGPTETADVTVTGLDLPPTGGRPV--PRWRY---VLDAGJRPVPGVGBLY-- 830
 Y 412 FYQAGPHLAAANPQTDGHTLQSQTGSSNTPESTYDYLJCGGAP-----LLAR 463
 Db 831 --LAGQYVARYG3YLGQ--HALTAER-----PVACPCKPGERMYRTGDLAR 871
 Y 464 L---IYLERD-----AGAF----TGGHGDAKYY 487
 Db 872 WLPDGHLVYVGGEDEQVKIRGFRIEPGEVEAALREDEGVAAVITYREDTPGTRLGVIV 931

862 TIAAVKRFTBEYAAI--NFINLPPCTIGELAQFYMANLLKYCDHSQVLINTLTSITG 919
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
12499 LLATGSCLAEVCVGRSLLEPTPT-----NMASATISHTRSA---- 12536
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
920 RRPRDPSSVLHWIRDVTSADIEOAKALLEKTEINLPFWTTAFTSTHVLRAAMNORPM 979
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
12537 ----PSDQV--VSPELQMASNVSSQQRSYSWRVTQVFPRPGILARPHPAVPCHELRSQ 12589
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
980 VV----LGISISKYHGAAGNNVFOAGNWGLNGKRNVCPLFTFDRTRRFIACPRG 1032
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
12590 CVQDAWPSAVGCSVR--SVRSTNSLMSFQRNLAGTS-----RPRRF---- 12631
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
1033 GFICPVTGPSSGNRNTTLDQVRGIVSGGAMVQLAIAYTVRAYGARAQHMAFDWL-SL 1092
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
12632 -----SERRRTTVIESTSPCTSG-----RRTCRAGASSISCAI---CSR 12670
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
1093 TDDEFIARDLDBLHDILQTLERPTNTVEGALEAVKILDEKTTAGDEGETTNLAFNFDSE 1152
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
12671 TSSS -----TRERSSGD-----R 12684
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
7 1153 PSHDITTSVNLTSGNSNISGS 1172
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
12685 PANDTISASASAPPGSSVSGA 12704
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

search completed: January 30, 2004, 13:20:11
db time: 49 secs

i903590KP1



lion date: 03-03-2004
king Officer: HABDULEKADER - HUDA ABDULEKADER
n: OIPEScanning
sier: 10190359

gual Date: 09-16-2002

o.	Doccode	Number of pages
1	NPL	12 /
2	NPL	4 /
3	NPL	5 /
4	NPL	8 /
5	NPL	9 /
6	NPL	4 /
7	NPL	5 /
8	NPL	2 /
9	NPL	13 /
10	NPL	6 /
11	NPL	14 /
12	NPL	6 /
13	NPL	7 /
14	NPL	8 /
15	NPL	40 /
16	NPL	6 /
17	NPL	7 /
18	NPL	8 /
19	NPL	4 /
20	NPL	3 /
21	NPL	4 /
22	NPL	30 /
23	NPL	8 /
24	NPL	4 /
25	NPL	8 /
26	NPL	17 /
27	NPL	14 /
28	NPL	5 /
29	NPL	9 /
30	NPL	12 /
31	NPL	7 /
		10 /